On the 3-Receiver Broadcast Channel with Degraded Message Sets and Confidential Messages

Li-Chia Choo, and Kai-Kit Wong, *Senior Member, IEEE*

Abstract

In this paper, bounds to the rate-equivocation region for the general 3-receiver broadcast channel (BC) with degraded message sets, are presented for confidential messages to be kept secret from one of the receivers. This model is more general than the 2-receiver BCs with confidential messages with an external wiretapper, and the recently studied 3-receiver degraded BCs with confidential messages, since in the model studied in this paper, the conditions on the receivers are general and the wiretapper receives the common message. Wyner's code partitioning combined with double-binning is used to show the achievable rate tuples. Error probability analysis and equivocation calculation are also provided. The secure coding scheme is sufficient to provide security for the 3-receiver BC with 2 or 3 degraded message sets, for the scenarios: (i) 3 degraded message sets, where the first confidential message is sent to receivers 1 and 2 and the second confidential message is sent to receiver 1, (ii) 2 degraded message sets, where one confidential message is sent to receiver 1, and (iii) 2 degraded message sets, where one confidential message is sent to receivers 1 and 2. The proof for the outer bound is shown for the cases where receiver 1 is more capable than the wiretap receiver 3, for the first two scenarios. Under the condition that both receivers 1 and 2 are less noisy than the wiretap receiver 3, the inner and outer bounds coincide, giving the rate-equivocation region for (iii). In addition, a new outer bound for the general 3-receiver BC with 3 degraded messages is obtained.

I. Introduction

Wireless communications channels today are vulnerable to eavesdropping or wiretapping due to the open nature of the channel, making the characterization of transmission rates for secure and reliable communication for the physical layer an important issue. In the wireless broadcast medium, the model of the broadcast channel (BC) with confidential messages, which was studied by Csiszár and Körner [1], is used to study simultaneously secure

The authors are with the Department of Electronic and Electrical Engineering, University College London, Torrington Place, London WC1E 7JE, United Kingdom (email: {l.choo, kwong}@ee.ucl.ac.uk).

The material in this paper will be presented in part at the International Conference on Wireless Communications and Signal Processing 2009, Nov. 13-15, Nanjing, China, 2009.

and reliable communication. The model in [1] is a generalization of the characterization of the wiretap channel by Wyner [2]. In [1], a common message is sent to 2 receivers, while a confidential message is sent to one of the receivers and kept secret from the other. The secrecy level is determined by the equivocation rate, which is the entropy rate of the confidential message conditioned on the channel output at the eavesdropper or wiretapper. The secrecy capacity region is defined as the set of transmission rates where the legitimate receiver decodes its confidential message while keeping the message secret from the wiretapper.

In more recent studies on the BC with confidential messages, Liu *et al.* [3] studied the scenario where there are 2 receivers and private messages are sent to each one and kept secret from the unintended receiver, while Xu *et al.* [4] looked at the same model in [3] but with a common message to both receivers. Then, Bagherikaram *et al.* [5] addressed the scenario where there are 2 receivers and one wiretapper, with confidential messages sent to the receivers. There have been recent studies where more than 2 receivers were considered. The authors in [6] and Ekrem and Ulukus in [7] independently studied the K-receiver BC with an external wiretapper. In [6], the K-receiver BC with confidential messages sent to each receiver was studied, while in [7], the same scenario was studied with the addition that each receiver also received a common message. Both used the degraded BC. In another work, an achievable inner bound for the K-receiver BC with a common message sent to all receivers and a confidential message sent to each of the receivers to be kept secret from an external wiretapper was derived by Kobayashi *et al.* in [8] for general conditions on the receivers' and wiretapper's channels. Finally, Chia and El Gamal in [9] derived an achievable inner bound for the 3-receiver BC with a common message sent to all receivers and a private message sent to 2 of the receivers to be kept secret from the third.

Recently in [10]–[12], Nair and El Gamal introduced the channel model of the 3-receiver BC with degraded message sets. In the general form of this model, a common message W_0 is sent to all of the receivers, denoted by the set $\mathbb{R}_{\text{all}}$, and the private messages, $W_i, W_{i-1}, \dots, W_1$, are sent to subsets of receivers $\mathbb{R}_i \subset \mathbb{R}_{i-1} \subset \cdots \subset \mathbb{R}_1 \subset \mathbb{R}_{\text{all}}$. This model best describes a multimedia broadcasting system, in which the common message W_0 may represent the lowest quality transmission, and W_1 the next higher quality transmission, and so on. In [10]–[12], three types of 3-receiver BCs with degraded message sets are studied:

1) 3-receiver BC with 3 degraded message sets where W_0 is sent to all three receivers, W_1 is sent to receivers 1 and 2, and a second private message W_2 is sent to receiver 1;
2) 3-receiver BC with 2 degraded message sets (Type 1) where the common message W_0 is sent to all three receivers and a private message W_1 is sent to the first receiver;
3) 3-receiver BC with 2 degraded message sets (Type 2) where the common message W_0 is sent to all three receivers and a private message W_1 sent to receivers 1 and 2.

While preparing this paper for submission, the authors became aware that Nair and El Gamal in [12] used a different coding scheme for their achievability proof compared to their earlier work [10], with detailed proofs in [11]. The added ingredient is rate splitting. However, a coding scheme with and without rate splitting is shown to give the same rate region in [12]. Based on this, in this paper, we shall not use rate splitting but base our achievability proof on the one in [10, 11].

The objective of this paper is to study this model of the 3-receiver BC with degraded message sets of [10], [11] with *secrecy* constraints. In particular, we characterize the transmission rates for the three types of 3-receiver BCs with degraded message sets from the model mentioned above where receiver 3 is a wiretapper. We note that the insights which this model of the 3-receiver BC with degraded message sets might bring are due to it being a more general model than the 2- or 3-receiver degraded BC with secrecy constraints. We also note that Chia and El Gamal in [9] have also studied the 3-receiver BC with 2 degraded message sets (Type 2) with receiver 3 being a wiretapper, but using a different coding scheme.

For the 3-receiver BC with 3 degraded message sets and 2 degraded message sets (Type 1) without secrecy constraints, the inner capacity bound in [10], [11] is achievable by superposition coding, Marton's achievability technique [13] and indirect decoding, where the receivers decoding the common message only do so via satellite codewords instead of cloud centers. For the general 3-receiver BC with degraded message sets, an outer bound to the capacity region was given in [10, 11] only for the general 3-receiver BC with 2 degraded message sets (Types 1 and 2). For the 3-receiver BC with 2 degraded message sets (Type 2), the inner and outer bounds coincide under the condition that first and second receivers are less noisy than the third receiver.

In our earlier work [14], we had studied the 3-receiver BC with 2 degraded message sets (Type 1), with the third receiver regarded as a wiretapper from which the private message is to be kept secret. In this paper, we consider the more general model of the 3-receiver BC with 3 degraded message sets where the third receiver is a wiretapper from which the private messages W_1, W_2 are to be kept secret. As the wiretapper in this case also decodes the common message, the 3-receiver BC with 3 degraded message sets with the third receiver a wiretapper describes a more general scenario than three types of scenarios: the 2-receiver BCs with an external wiretapper of [5], the 2-receiver BC with 3 degraded message sets and an external wiretapper, and the 3-receiver degraded BCs with an external wiretapper by the virtue of the general conditions on the receivers.

In our secure coding scheme, we shall use a combination of the code partitioning of Wyner [2] and double-binning of Liu *et al.* [3] to show the achievability of an inner bound to the rate-equivocation region for the 3-receiver BC with 3 degraded message sets. Error probability analysis and equivocation calculation for the private messages are provided. The proposed secure coding scheme is shown to be sufficient for providing security for both the 3-receiver BC with 3 degraded message sets and the 3-receiver BC with 2 degraded message sets (Type 1). We obtain outer bounds to the rate-equivocation region for the 3-receiver BC with 3 degraded message sets for the case where receiver 1 is more capable than the wiretap receiver 3, a weaker condition than the condition that receiver 3 is a degraded version of receiver 1 or the condition that receiver 1 is less noisy than the wiretap receiver 3 [15]. By removing the security constraints, we further obtain an outer bound to the capacity region for the *general* 3-receiver BC with 3 degraded message sets, which is not found in [10]– [12]. This is because the condition that receiver 1 is more capable than receiver 3 applies only to the case where we have secrecy constraints. Then, we show that the outer bounds to the rate-equivocation region for the 3-receiver BC with 3 degraded message sets reduce to the outer bounds to the rate-equivocation region for the 3-receiver BC with 2 degraded message sets (Type 1), if receiver 1 is more capable than the wiretap receiver 3. Finally, we show that, under the condition that the first and second receivers are less noisy than the third receiver, respectively (still a more general condition than degradedness [15]), the inner and outer bounds to the rate-equivocation region for the 3-receiver BC with 3 degraded message sets reduce to the region for the 3-receiver BC with 2 degraded message sets (Type 2). This rate-equivocation region we obtain is furthermore a special case of the variant of the 3-receiver BC with 2 degraded message sets (Type 2) studied in [9] with a different coding scheme.

This paper is organized as follows. In Section II, we describe the model for the 3-receiver BC with degraded message sets. In Section III, we state our main results, the bounds to the rate-equivocation region. In Section IV, we show achievability of the inner bound to the rate-equivocation region using our secure coding scheme for the 3-receiver BC with 3 degraded message sets and the 3-receiver BC with 2 degraded message sets (Type 1) and show error probability analysis and equivocation calculation for the private messages. We show that the coding scheme provides security for both types of channel. In Section V, we show the proof of the outer bounds for the three types of the 3-receiver BC with degraded message sets. Lastly, we give conclusions in Section VI.

II. The 3-Receiver BC with Degraded Message Sets

In this paper, we use the uppercase letter to denote a random variable (e.g., X) and the lowercase letter for its realization (e.g., x). The alphabet set of X is denoted by $\mathcal{X}$ so that $x \in \mathcal{X}$. We denote a sequence of n random variables by $\mathbf{X} = (X_1, \ldots, X_n)$ with its realization $\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{X}^n$ if $x_i \in \mathcal{X}$ for $i = 1, 2, \ldots, n$. Furthermore, we define the subsequences of $\mathbf{X}$ as $\mathbf{X}^i \triangleq (X_1, X_2, \ldots, X_i)$ and $\tilde{\mathbf{X}}^i \triangleq (X_i, \ldots, X_n)$.

The discrete memoryless BC with 3 receivers has an input random sequence, $\mathbf{X}$, and 3 output random sequences at the receivers, denoted respectively by $\mathbf{Y}_1, \mathbf{Y}_2$ and $\mathbf{Y}_3$, all of length n, with $\mathbf{x} \in \mathcal{X}^n$, $\mathbf{y}_1 \in \mathcal{Y}_1^n$, $\mathbf{y}_2 \in \mathcal{Y}_2^n$, and $\mathbf{y}_3 \in \mathcal{Y}_3^n$. The conditional distribution for n uses of the channel is given by

$$p(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3|\mathbf{x}) = \prod_{i=1}^{n} p(y_{1i}, y_{2i}, y_{3i}|x_i). \tag{1}$$

A $(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, n)$-code for the 3-receiver BC with 3 degraded message sets, as depicted in Figure 1,

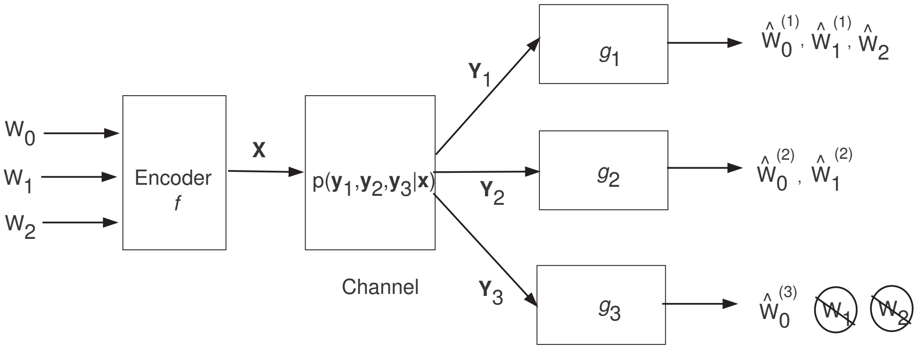


Fig. 1. The 3-receiver BC with 3 degraded message sets and confidential messages.

consists of the following parameters:

$$\mathcal{W}_0 = \left\{1, \ldots, 2^{nR_0}\right\}, \text{(common message set)}$$

$$\mathcal{W}_1 = \left\{1, \ldots, 2^{nR_1}\right\}, \text{(private message set)},$$

$$\mathcal{W}_2 = \left\{1, \ldots, 2^{nR_2}\right\}, \text{(private message set)},$$

$$f : \mathcal{W}_0 \times \mathcal{W}_1 \times \mathcal{W}_2 \mapsto \mathcal{X}^n, \text{(encoding function)},$$

$$g_1 : \mathcal{Y}_1^n \mapsto \mathcal{W}_0 \times \mathcal{W}_1 \times \mathcal{W}_2, \text{(decoding function 1)},$$

$$g_2 : \mathcal{Y}_2^n \mapsto \mathcal{W}_0 \times \mathcal{W}_1, \text{(decoding function 2)},$$

$$g_3 : \mathcal{Y}_3^n \mapsto \mathcal{W}_0, \text{(decoding function 3)}.$$

In particular, we have $g_1(\mathbf{Y}_1) = (\hat{W}_0^{(1)}, \hat{W}_1^{(1)}, \hat{W}_2)$, $g_2(\mathbf{Y}_2) = (\hat{W}_0^{(2)}, \hat{W}_1^{(2)})$, and $g_3(\mathbf{Y}_3) = \hat{W}_0^{(3)}$, where the notation "$\hat{(\cdot)}$" highlights that the decoded messages are estimates, with the error probability

$$P_e^{(n)} = \Pr\left\{(\hat{W}_0^{(1)}, \hat{W}_0^{(2)}, \hat{W}_0^{(3)}, \hat{W}_1^{(1)}, \hat{W}_1^{(2)}, \hat{W}_2) \neq (W_0, W_0, W_0, W_1, W_1, W_2)\right\}. \tag{2}$$

In this setup, Y_3 is the wiretapper, and the secrecy level of the messages sent are as follows:

1) For W_1 sent to users 1 and 2, the secrecy level is defined by the equivocation rate $\frac{1}{n}H(W_1|\mathbf{Y}_3)$;
2) For W_2 sent to user 1, the secrecy level is defined by the equivocation rate $\frac{1}{n}H(W_2|\mathbf{Y}_3)$;
3) The combined message (W_1, W_2) sent to user 1 has secrecy level defined by the equivocation rate $\frac{1}{n}H(W_1, W_2|\mathbf{Y}_3)$.

In addition, a $(2^{nR_0}, 2^{nR_1}, n)$-code for the 3-receiver BC with 2 degraded message sets (Type 1), as shown in

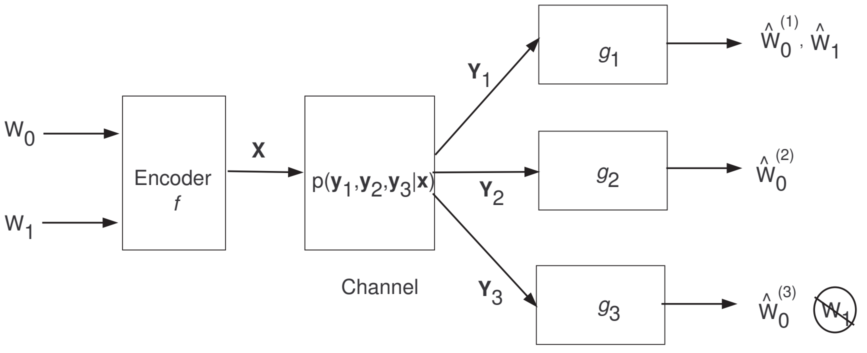


Fig. 2. The 3-receiver BC with 2 degraded message sets (Type 1) and confidential message.

Figure 2, consists of the following parameters:

$$\mathcal{W}_0 = \left\{1, \ldots, 2^{nR_0}\right\}, \text{(common message set)}$$

$$\mathcal{W}_1 = \left\{1, \ldots, 2^{nR_1}\right\}, \text{(private message set)},$$

$$f : \mathcal{W}_0 \times \mathcal{W}_1 \mapsto \mathcal{X}^n, \text{(encoding function)},$$

$$g_1 : \mathcal{Y}_1^n \mapsto \mathcal{W}_0 \times \mathcal{W}_1, \text{(decoding function 1)},$$

$$g_2 : \mathcal{Y}_2^n \mapsto \mathcal{W}_0, \text{(decoding function 2)},$$

$$g_3 : \mathcal{Y}_3^n \mapsto \mathcal{W}_0, \text{(decoding function 3)}.$$

We have $g_1(\mathbf{Y}_1) = (\hat{W}_0^{(1)}, \hat{W}_1^{(1)})$, $g_2(\mathbf{Y}_2) = \hat{W}_0^{(2)}$, and $g_3(\mathbf{Y}_3) = \hat{W}_0^{(3)}$, with the error probability

$$P_e^{(n)} = \Pr\left\{(\hat{W}_0^{(1)}, \hat{W}_0^{(2)}, \hat{W}_0^{(3)}, \hat{W}_1^{(1)}) \neq (W_0, W_0, W_0, W_1,)\right\}. \tag{3}$$

With Y_3 the wiretapper, and the secrecy level of the message sent is $\frac{1}{n}H(W_1|\mathbf{Y}_3)$.

Finally, a $(2^{nR_0}, 2^{nR_1}, n)$-code for the 3-receiver BC with 2 degraded message sets (Type 2), as shown in Figure 3, consists of the parameters:

$$\mathcal{W}_0 = \left\{1, \ldots, 2^{nR_0}\right\}, \text{(common message set)}$$

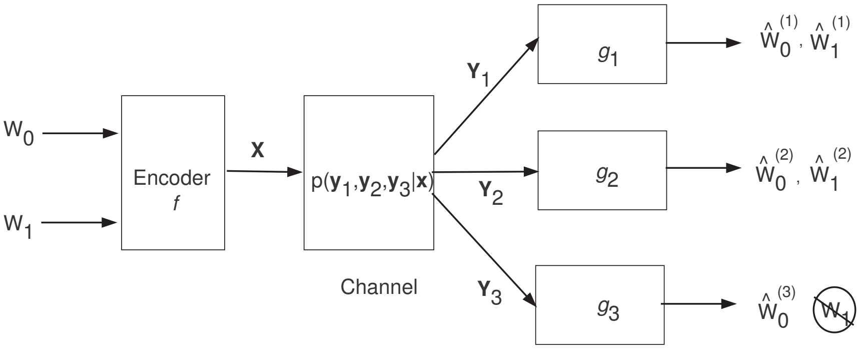
$$\mathcal{W}_1 = \left\{1, \ldots, 2^{nR_1}\right\}, \text{(private message set)},$$

$$f : \mathcal{W}_0 \times \mathcal{W}_1 \mapsto \mathcal{X}^n, \text{(encoding function)},$$

$$g_1 : \mathcal{Y}_1^n \mapsto \mathcal{W}_0 \times \mathcal{W}_1, \text{(decoding function 1)},$$

$$g_2 : \mathcal{Y}_2^n \mapsto \mathcal{W}_0 \times \mathcal{W}_1, \text{(decoding function 2)},$$

$$g_3 : \mathcal{Y}_3^n \mapsto \mathcal{W}_0, \text{(decoding function 3)}.$$



W_0
W_1
Encoder
f
$\mathbf{X}$
p($\mathbf{y}_1$,$\mathbf{y}_2$,$\mathbf{y}_3$|$\mathbf{x}$)
Channel
$\mathbf{Y}_1$
$\mathbf{Y}_2$
$\mathbf{Y}_3$
g_1
g_2
g_3
$\hat{W}_0^{(1)}$, $\hat{W}_1^{(1)}$
$\hat{W}_0^{(2)}$, $\hat{W}_1^{(2)}$
$\hat{W}_0^{(3)}$
W_1

Fig. 3. The 3-receiver BC with 2 degraded message sets (Type 2) and confidential message.

We have $g_1(\mathbf{Y}_1) = (\hat{W}_0^{(1)}, \hat{W}_1^{(1)})$, $g_2(\mathbf{Y}_2) = (\hat{W}_0^{(2)}, \hat{W}_1^{(2)})$, and $g_3(\mathbf{Y}_3) = \hat{W}_0^{(3)}$, and error probability

$$P_e^{(n)} = \Pr\left\{(\hat{W}_0^{(1)}, \hat{W}_0^{(2)}, \hat{W}_0^{(3)}, \hat{W}_1^{(1)}, \hat{W}_1^{(2)}) \neq (W_0, W_0, W_0, W_1, W_1)\right\}. \tag{4}$$

The secrecy level of the message W_1 sent to users 1 and 2 is defined by the equivocation rate $\frac{1}{n}H(W_1|\mathbf{Y}_3)$.

III. Bounds to the Rate-Equivocation Region

A. The 3-Receiver BC with 3 Degraded Message Sets

For the 3-receiver BC with 3 degraded message sets, the rate tuple $(R_0, R_1, R_{1e}, R_2, R_{2e})$ is said to be achievable if for any $\eta, \epsilon_1, \tilde{\epsilon}_1, \epsilon_2, \epsilon_{1,2} > 0$, there exists a sequence of $(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, n)$-codes for which $P_e^{(n)} \leq \eta$ and the equivocation rates R_{1e} and R_{2e} satisfy

$$\frac{1}{n}H(W_1|\mathbf{Y}_3) \geq R_{1e} - \epsilon_1, \text{ or } \frac{1}{n}H(W_1|\mathbf{Y}_3) \geq R_{1e} - \tilde{\epsilon}_1, \tag{5a}$$

$$\frac{1}{n}H(W_2|\mathbf{Y}_3) \geq R_{2e} - \epsilon_2, \tag{5b}$$

$$\frac{1}{n}H(W_1, W_2|\mathbf{Y}_3) \geq R_{1e} + R_{2e} - \epsilon_{1,2}. \tag{5c}$$

The two conditions on W_1 arise because the equivocation rate depends on which destination W_1 is sent to, as can be seen below in (6d). Recall from the model of the 3-receiver BC with 3 degraded message sets that W_1 is sent to both Y_1 and Y_2. The first equivocation rate in (5a) corresponds to W_1 being sent to receiver Y_2 and the second equivocation rate in (5a) corresponds to W_1 being sent to receiver Y_1. The rate-equivocation region for the 3-receiver BC with 3 degraded message sets is the closure of the set of all rate-tuples such that $(R_0, R_1, R_{1e}, R_2, R_{2e})$ is achievable. Our analysis does not include the case of perfect secrecy (i.e., the rate region with $R_{1e} = R_1$ and $R_{2e} = R_2$). The following theorems summarize the main results of this paper.

Theorem 1: An inner bound to the rate-equivocation region for the 3-receiver BC with 3 degraded message sets is the closure of all rate-tuples $(R_0, R_1, R_{1e}, R_2, R_{2e})$ satisfying

$$\begin{aligned}
R_{1e} \leq& R_1, && \text{(6a)}\\
R_{2e} \leq& R_2, && \text{(6b)}\\
R_0 \leq& I(U_3; Y_3), && \text{(6c)}\\
R_{1e} \leq& \min\left\{I(U_2; Y_2|U_1) - R_1', I(X; Y_1|U_3) - R_1' - R_2'\right\}, && \text{(6d)}\\
R_{2e} \leq& I(X; Y_1|U_2) - R_2', && \text{(6e)}\\
R_{1e} + R_{2e} \leq& I(X; Y_1|U_1) - R_1' - R_2', && \text{(6f)}\\
R_0 + R_1 \leq& \min\left\{I(U_2; Y_2), I(U_3; Y_3) + I(U_2; Y_2|U_1) - I(U_2; U_3|U_1)\right\} && \text{(6g)}\\
2R_0 + R_1 \leq& I(U_3; Y_3) + I(U_2; Y_2) - I(U_2; U_3|U_1), && \text{(6h)}\\
R_0 + R_2 \leq& I(U_3; Y_3) + I(X; Y_1|U_2, U_3), && \text{(6i)}\\
R_0 + R_1 + R_2 \leq& \min\left\{I(U_3; Y_3) + I(X; Y_1|U_3), I(X; Y_1),\right.\\
& \left. I(U_3; Y_3) + I(U_2; Y_2|U_1) - I(U_2; U_3|U_1) + I(X; Y_1|U_2, U_3)\right\}, && \text{(6j)}\\
2R_0 + R_1 + R_2 \leq& I(U_3; Y_3) + I(U_2; Y_2) - I(U_2; U_3|U_1) + I(X; Y_1|U_2, U_3), && \text{(6k)}\\
R_0 + 2R_1 + R_2 \leq& I(U_3; Y_3) + I(U_2; Y_2|U_1) - I(U_2; U_3|U_1) + I(X; Y_1|U_3), && \text{(6l)}\\
2R_0 + 2R_1 + R_2 \leq& I(U_3; Y_3) + I(U_2; Y_2) - I(U_2; U_3|U_1) + I(X; Y_1|U_3), && \text{(6m)}
\end{aligned}$$

in which $R_1' \triangleq I(U_2; Y_3|U_1)$ and $R_2' \triangleq I(X; Y_3|U_2)$ are defined over the the probability density function (p.d.f.)

$$p(u_1, u_2, u_3, x) = p(u_1)p(u_2|u_1)p(x, u_3|u_2) = p(u_1)p(u_3|u_1)p(x, u_2|u_3) = p(u_1)p(u_2, u_3|u_1)p(x|u_2, u_3), \quad (7)$$

which is induced by the coding scheme. In addition, we require that the condition

$$I(X; Y_3|U_2) \leq I(X; Y_1|U_2, U_3) \quad (8)$$

is met. From the p.d.f. (7), the auxiliary random variables U_1, U_2 and U_3 satisfy the Markov chain conditions

$$\begin{aligned}
&U_1 \to U_2 \to (U_3, X) \to (Y_1, Y_2, Y_3), && \text{(9a)}\\
&U_1 \to U_3 \to (U_2, X) \to (Y_1, Y_2, Y_3), && \text{(9b)}\\
&U_1 \to (U_2, U_3) \to X \to (Y_1, Y_2, Y_3). && \text{(9c)}
\end{aligned}$$

Proof: The proof of achievability is based on that for the 3-receiver BC with 3 degraded message sets in [10], [11] which uses Marton's achievability scheme [13] combined with superposition coding and is given in Section IV-A with the equivocation calculation (bounds for R_{e1}, R_{e2}) to be presented in Section IV-C. ■

Since our achievability scheme is based upon that of [10], [11], it is natural that the inner bound is the same as that of [10], [11], but with the addition of the equivocation rates. In fact it will be the same as [12], with the addition of the equivocation rates. As a check, setting $Y_1 = Y_3$ in (6d)–(6f), $R_{1e} \leq 0$, $R_{2e} \leq 0$ and $R_{1e}+R_{2e} \leq 0$, so no secrecy rate is possible. Thus the equivocation rates (6d)–(6f) are achievable.

Theorem 2: An outer bound to the rate-equivocation region for the 3-receiver BC with 3 degraded message sets, where Y_1 is more capable than Y_3, is the closure of all rate-tuples $(R_0, R_1, R_{1e}, R_2, R_{2e})$ that satisfies

$$R_{1e} \leq R_1, \tag{10a}$$

$$R_{2e} \leq R_2, \tag{10b}$$

$$R_0 \leq \min\left\{I(U_1;Y_1), I(U_3;Y_3) - I(U_3;Y_1|U_1)\right\}, \tag{10c}$$

$$R_{1e} \leq \min\left\{I(U_2;Y_2|U_1) - I(U_2;Y_3|U_1), I(X;Y_1|U_3) - I(X;Y_3|U_1)\right\}, \tag{10d}$$

$$R_{2e} \leq I(X;Y_1|U_2) - I(X;Y_3|U_2), \tag{10e}$$

$$R_{1e} + R_{2e} \leq I(X;Y_1|U_1) - I(X;Y_3|U_1), \tag{10f}$$

$$\begin{aligned} R_0 + R_1 \leq \min\{&I(U_2;Y_1), I(U_2;Y_2), I(U_1;Y_1) + I(U_2;Y_2|U_1), \\ &I(U_3;Y_3) + I(U_2;Y_1|U_1), I(U_3;Y_3) + I(U_2;Y_2|U_1)\}, \end{aligned} \tag{10g}$$

$$R_0 + R_2 \leq \min\left\{I(U_1;Y_1) + I(X;Y_1|U_2,U_3), I(U_3;Y_3) + I(X;Y_1|U_2,U_3)\right\}, \tag{10h}$$

$$\begin{aligned} R_0 + R_1 + R_2 \leq \min\{&I(X;Y_1), I(U_3;Y_3) + I(X;Y_1|U_3), \\ &I(U_1;Y_1) + I(U_2;Y_2|U_1) + I(X;Y_1|U_2,U_3), \\ &I(U_3;Y_3) + I(U_2;Y_2|U_1) + I(X;Y_1|U_2,U_3), I(U_2;Y_2) + I(X;Y_1|U_2,U_3)\}. \end{aligned} \tag{10i}$$

Proof: The proof for this outer bound is given in Section V-A. ■

We see that the equivocation rates for (R_{1e}, R_{2e}) in the inner and outer bounds in Theorems 1 and 2 match. Note that the equivocation rate for R_{1e} received at Y_1 is reduced by $\Delta_1 = I(U_2;Y_3|U_1)+I(X;Y_3|U_2) = I(X;Y_3|U_1)$. In Δ_1, the first term is needed to protect the codewords generated by Marton's achievability scheme, and the second term protects codewords generated by superposition coding. While it is only required to protect the codewords generated by Marton's achievability scheme for the general 2-receiver BC in [5], our secure scheme (to be presented in Section IV) does this, as well as protects the additional codewords generated by superposition coding. Hence, our secure scheme results in a loss for R_{1e} (compared to R_1) that may be larger than expected.

It is also noted that by removing the secrecy constraints from the outer bound to the rate-equivocation region for the 3-receiver BC with 3 degraded message sets, we can obtain a new outer bound to the capacity region of the *general* 3-receiver BC with 3 degraded message sets without secrecy. We see this by setting $R_{1e} = 0$ and

$R_{2e} = 0$ in Theorem 2 above. Since the restriction that receiver Y_1 is more capable than receiver Y_3 is only applicable when deriving R_{1e} and R_{2e} as will be shown in Section V-A, removing the secrecy constraints will give us the outer bound to the capacity region of the general 3-receiver BC with 3 degraded message sets.

Theorem 3: An outer bound to the capacity region for the general 3-receiver BC with 3 degraded message sets is the closure of all rate-tuples (R_0, R_1, R_2) satisfying

$$R_0 \leq \min\left\{I(U_1;Y_1), I(U_3;Y_3) - I(U_3;Y_1|U_1)\right\}, \tag{11a}$$

$$\begin{aligned} R_0 + R_1 \leq \min\{&I(U_2;Y_1), I(U_2;Y_2), I(U_1;Y_1) + I(U_2;Y_2|U_1), \\ &I(U_3;Y_3) + I(U_2;Y_1|U_1), I(U_3;Y_3) + I(U_2;Y_2|U_1)\}, \end{aligned} \tag{11b}$$

$$R_0 + R_2 \leq \min\left\{I(U_1;Y_1) + I(X;Y_1|U_2,U_3), I(U_3;Y_3) + I(X;Y_1|U_2,U_3)\right\}, \tag{11c}$$

$$\begin{aligned} R_0 + R_1 + R_2 \leq \min\{&I(X;Y_1), I(U_3;Y_3) + I(X;Y_1|U_3), I(U_1;Y_1) + I(U_2;Y_2|U_1) + I(X;Y_1|U_2,U_3), \\ &I(U_3;Y_3) + I(U_2;Y_2|U_1) + I(X;Y_1|U_2,U_3), \, I(U_2;Y_2) + I(X;Y_1|U_2,U_3)\}. \end{aligned} \tag{11d}$$

Proof: As described above. ■

B. The 3-Receiver BC with 2 Degraded Message Sets

The 3-receiver BC with 3 degraded message sets with secrecy constraints can be specialized to 2 classes of a 3-receiver BC with 2 degraded message sets with secrecy constraints:

1) Type 1: A 3-receiver BC where (W_0, W_1) is sent to receiver Y_1 and W_0 is sent to receivers Y_2 and Y_3, where W_1 is to be kept secret from receiver Y_3;
2) Type 2: A 3-receiver BC where (W_0, W_1) is sent to receivers Y_1 and Y_2 and W_0 is sent to receiver Y_3, where W_1 is to be kept secret from receiver Y_3.

We note that the inner and outer bounds do not match for the first case, but match for the second case under the condition that both receivers Y_1 and Y_2 are less noisy than receiver Y_3.

We have studied the Type 1 channel in [14]. In this paper, we shall briefly review the achievability scheme for secrecy constraints to see the differences from the 3 degraded message sets case, and show that the outer bound for the 3 degraded message sets case can be reduced to the outer bound for this Type 1 channel.

For the Type 2 channel, we shall show that the bounds on the rate-equivocation region can be specialized from the 3 degraded message sets case. We also note that the Type 2 channel is a special case of the inner bound to the rate-equivocation region for a 3-receiver BC with 2 degraded message sets studied in Chia and El Gamal [9] using a different coding scheme. In [9], the message reception and secrecy conditions are the same as the Type 2 channel. Thus, both our bounds and that of [9] will reduce to the Type 2 channel. Also, our outer bounds will reduce to the Type 2 channel under the conditions that both receivers Y_1 and Y_2 are less noisy than receiver Y_3.

We state the inner and outer bounds to the rate-equivocation region for the Type 1 channel in Corollaries 1 and 2, and the rate-equivocation region for the Type 2 channel in Corollary 3.

Corollary 1: An inner bound to the rate-equivocation region for the 3-receiver BC with 2 degraded message sets (Type 1) is the closure of all rate-tuples (R_0, R_1, R_{1e}) satisfying

$$R_{1e} \leq R_1 \tag{12a}$$

$$R_0 \leq \min\{I(U_2;Y_2), I(U_3;Y_3)\} \tag{12b}$$

$$R_{1e} \leq \min\{I(X;Y_1|U_1) - \Delta_2, I(X;Y_1|U_2) + I(X;Y_1|U_3) - I(X;Y_3|U_2) - \Delta_2\}, \tag{12c}$$

$$2R_0 \leq I(U_2;Y_2) + I(U_3;Y_3) - I(U_2;U_3|U_1) \tag{12d}$$

$$R_0 + R_1 \leq \min\{I(X;Y_1), I(U_2;Y_2) + I(X;Y_1|U_2), I(U_3;Y_3) + I(X;Y_1|U_3)\}, \tag{12e}$$

$$2R_0 + R_1 \leq I(U_2;Y_2) + I(U_3;Y_3) - I(U_2;U_3|U_1) + I(X;Y_1|U_2,U_3), \tag{12f}$$

$$2R_0 + 2R_1 \leq I(U_2;Y_2) + I(X;Y_1|U_2) + I(U_3;Y_3) + I(X;Y_1|U_3) - I(U_2;U_3|U_1), \tag{12g}$$

under the same Markov chain conditions (9) for the auxiliary random variables, where $\Delta_2 \triangleq I(U_2;Y_3|U_1) + I(X;Y_3|U_2)$, and the conditions

$$\begin{cases} I(X;Y_3|U_2) \leq I(X;Y_1|U_2,U_3), \\ I(X;Y_3|U_2) \leq I(X;Y_1|U_2), \end{cases} \tag{13}$$

are satisfied.

Proof: See Section IV-B for the achievability proof, and [14] for the equivocation calculation. ■

We see that Δ_2 in Corollary 2 may be expressed as

$$\Delta_2 \triangleq I(U_2;Y_3|U_1) + I(X;Y_3|U_2) = I(X;Y_3|U_1), \tag{14}$$

which is $\geq I(X;Y_3|U_3)$. Thus, as a check, when $Y_1 = Y_3$ in (12c), $R_{1e} \leq 0$, so no secrecy rate is possible and therefore the equivocation rate (12c) is achievable. Also, when compared to the equivocation rates on R_{1e} for the 3 degraded message sets channel in (6d), a smaller rate is achievable for W_1 sent to Y_1. Then, by the virtue of sending W_2 to Y_1, the coding scheme of [10], [11] is able to give a higher equivocation rate for W_1 sent to Y_1. It appears that by sending more messages to receiver Y_1, then the achievable equivocation rates can be increased.

The lower achievable rate for W_1 sent to Y_1 for the 2 degraded message sets (Type 1) channel is due to the fact that the achievable coding scheme protects all the codewords generated by superposition coding. We note that the coding scheme of [10, 11] generates codewords giving rise to the rates $R_1 \leq I(X;Y_1|U_2) + I(X;Y_1|U_3)$ and $R_1 \leq I(X;Y_1|U_1)$. From the fact that when $Y_1 = Y_3$ in (12c), $R_{1e} \leq 0$ for both choices of R_{1e}, so implying the equivocation rates (12c) are achievable, we see that our proposed secure scheme is able to protect all the codewords generated by superposition coding, but with a smaller achievable equivocation rate for W_1 sent to Y_1 compared to R_{1e} (with W_1 sent to Y_1) for the 3 degraded message sets channel.

The outer bound for the Type 1 3-receiver 2 degraded message sets BC is stated as follows.

Corollary 2: An outer bound to the rate-equivocation region for the 3-receiver BC with 2 degraded message sets (Type 1), where Y_1 is more capable than Y_3, is the closure of all rate-tuples (R_0, R_1, R_{1e}) satisfying

$$R_{1e} \leq R_1, \tag{15a}$$

$$R_0 \leq \min\{I(U_1;Y_1), I(U_2;Y_2) - I(U_2;Y_1|U_1), I(U_3;Y_3) - I(U_3;Y_1|U_1)\} \tag{15b}$$

$$R_{1e} \leq I(X;Y_1|U_1) - I(X;Y_3|U_1), \tag{15c}$$

$$R_0 + R_1 \leq \min\{I(X;Y_1), I(U_2;Y_2) + I(X;Y_1|U_2), I(U_3;Y_3) + I(X;Y_1|U_3)\}. \tag{15d}$$

Proof: See Section V-B. ■

We state the rate-equivocation region for the Type 2 3-receiver 2 degraded message sets BC below.

Corollary 3: The secrecy capacity region for the 3-receiver BC with 2 degraded message sets (Type 2) for the case where Y_1 and Y_2 are both less noisy than Y_3 is the closure of all rate-tuples (R_0, R_1, R_{1e}) satisfying

$$R_{1e} \leq R_1, \tag{16a}$$

$$R_0 \leq I(U;Y_3), \tag{16b}$$

$$R_{1e} \leq \min\{I(X;Y_1|U) - I(X;Y_3|U), I(X;Y_2|U) - I(X;Y_3|U)\}, \tag{16c}$$

$$R_0 + R_1 \leq \min\{I(X;Y_1), I(X;Y_2)\}, \tag{16d}$$

over the p.d.f. $p(u,x) = p(u)p(x|u)$.

Proof: In this channel class, the inner and outer bounds match. The proof of achievability follows by using code partitioning for security, as in [1, 2], where it can be seen that the codeword $\mathbf{X}$ is protected by the partition of $I(X;Y_3|U)$. The rate-equivocation region is achievable by setting $R_2 = 0$, $R_{2e} = 0$, $U_2 = X$, $U_3 = U_1 = U$ in Theorems 1 and 2 and using the conditions that Y_1 and Y_2 are less noisy than Y_3. Therefore, we have the conditions $I(U;Y_3) \leq I(U;Y_1)$ and $I(U;Y_3) \leq I(U;Y_2)$. See Section V-C for the converse proof. ■

It is worth emphasizing here that this channel class is more general than the special case of the 3-receiver BC with 2 degraded message sets (Type 1) under the condition that Y_1 is less noisy than Y_3 in [14], since Y_2 receives W_1 here but this is not the case in [14].

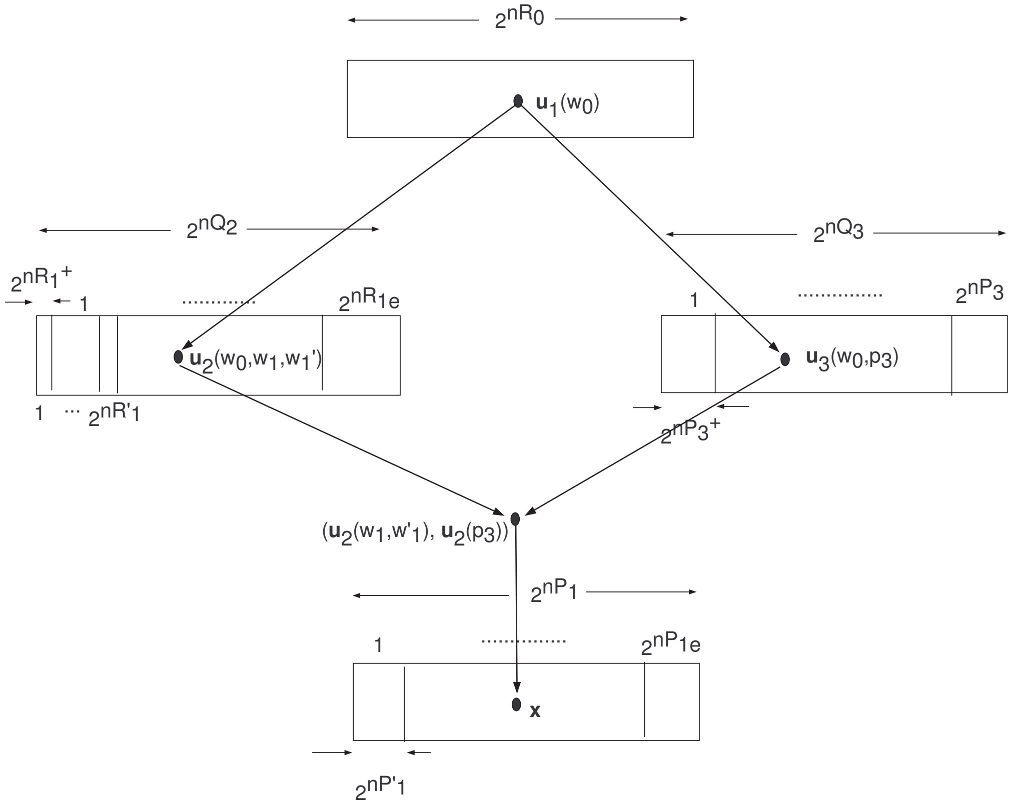


Fig. 4. Coding for 3-receiver BC with degraded message sets and confidential messages.

IV. Inner Bound for the 3-Receiver BC with 3 Degraded Message Sets

A. *Proof of Achievability for 3-Receiver BC with 3 Degraded Message Sets*

Our achievability proof for the 3-receiver BC with 3 degraded message sets is an alternative version of the one in [11, Appendix III]. We use Wyner's code partitioning [2] with the double-binning scheme of [3] to provide secrecy, together with the coding scheme for the 3-receiver BC with 3 degraded message sets in [10, 11].

The scheme of [10], [11] represents W_0 by U_1, then breaks W_2 into 2 parts. The first part is combined with U_1 by superposition coding to generate U_3. The message W_1 is combined with U_1 by superposition coding to generate U_2. U_2 and U_3 are partitioned into bins and the product bin containing the joint typical pair (achievable by Marton's coding scheme) is combined with the second part of W_2 by superposition coding to obtain X.

At the receivers, Y_1 decodes U_1, U_2, U_3, and X to recover the messages W_0, W_1 and W_2, while Y_2 decodes U_1 and U_2 to recover messages W_0 and W_1 and Y_3 decodes U_1 indirectly using U_3 to recover W_0. In our secure scheme, the codewords $\mathbf{U}_2$ and $\mathbf{X}$ are, respectively, protected from receiver Y_3 (i.e., the wiretapper) by a one-sided double-binning and code partitioning. This is depicted in Figure 4.

Suppose that we have the p.d.f. in (7) which induces the Markov chain conditions $U_1 \to U_2 \to (U_3, X)$, $U_1 \to U_3 \to (U_2, X)$ and $U_1 \to (U_2, U_3) \to X$. The following describes the encoding and decoding processes.

Codebook generation: Let $\tilde{R}_1 = R_{1e} + R'_1 + R^{\dagger}_1$, $\tilde{P}_3 = P_3 + P^{\dagger}_3$, $P_1 = P_{1e} + P'_1$, and $R_{2e} = P_3 + P_{1e}$. Define, for security,

$$P'_1 \triangleq I(X; Y_3|U_2) - \delta_1, \text{ and } R'_1 \triangleq I(U_2; Y_3|U_1) - \delta_1, \tag{17}$$

where $\delta_1 > 0$ and is small for n sufficiently large.

First of all, generate 2^{nR_0} sequences $\mathbf{U}_1(w_0)$, for $w_0 \in \mathcal{W}_0$, randomly and uniformly from the set of typical $\mathbf{U}_1$ sequences. For each $\mathbf{U}_1(w_0)$, generate 2^{nQ_2} sequences $\mathbf{U}_2(w_0, q_2)$ randomly and uniformly from the set of conditionally typical $\mathbf{U}_2$ sequences, and also 2^{nQ_3} sequences $\mathbf{U}_3(w_0, q_3)$ randomly and uniformly from the set of conditionally typical $\mathbf{U}_3$ sequences. Next, randomly partition the sequences, $\mathbf{U}_2(w_0, q_2)$, into $2^{n\tilde{R}_1}$ equally-sized bins, and the sequences, $\mathbf{U}_3(w_0, q_3)$, into $2^{n\tilde{P}_3}$ equally-sized bins. The $\mathbf{U}_2$ codewords undergo a double partition: the first into $2^{nR_{1e}}$ bins, and the second further partitions them into $2^{nR'_1}$ bins, each of size $2^{nR^{\dagger}_1}$. On the other hand, the $\mathbf{U}_3$ codewords undergo a single partition into 2^{nP_3} bins, each of size $2^{nP^{\dagger}_3}$.

Each product bin (w_1, w'_1, p_3) contains the joint typical pair $(\mathbf{U}_2(w_0, w_1, w'_1, w^{\dagger}_1), \mathbf{U}_3(w_0, p_3, p^{\dagger}_3))$ for $w_1 \in \{1, \ldots, 2^{nR_{1e}}\}$, $w'_1 \in \{1, \ldots, 2^{R'_1}\}$, $w^{\dagger}_1 \in \{1, \ldots, 2^{nR^{\dagger}_1}\}$, $p_3 \in \{1, \ldots, 2^{nP_3}\}$, and $p^{\dagger}_3 \in \{1, \ldots, 2^{nP^{\dagger}_3}\}$ with high probability under the conditions [16]

$$\begin{aligned} R_{1e} + R'_1 + R^{\dagger}_1 \le Q_2 &\Rightarrow R_{1e} + R'_1 \le Q_2, \\ P_3 + P^{\dagger}_3 \le Q_3 &\Rightarrow P_3 \le Q_3, \\ R^{\dagger}_1 + P^{\dagger}_3 &> I(U_2; U_3|U_1), \\ R_{1e} + R'_1 + P_3 &\le Q_2 + Q_3 - I(U_2; U_3|U_1). \end{aligned} \tag{18}$$

Now let us rewrite the joint typical pair as $(\mathbf{u}_2(w_0, w_1, w'_1), \mathbf{u}_3(w_0, p_3))$. For each such pair corresponding to the product bin (w_1, w'_1, p_3), generate 2^{nP_1} sequences of codewords $\mathbf{X}(w_0, w_1, w'_1, p_3, p_1, p'_1)$, for $p_1 \in \{1, \ldots, 2^{nP_{1e}}\}$ and $p'_1 \in \{1, \ldots, 2^{nP'_1}\}$, uniformly and randomly over the set of conditionally typical $\mathbf{X}$ sequences. The 2^{nP_1} codewords are partitioned into $2^{nP_{1e}}$ subcodes with $2^{nP'_1}$ codewords within the subcodes.

Encoding: To send (w_0, w_1, w_2), express w_2 by (p_1, p_3) and send the codeword $\mathbf{x}(w_0, w_1, w'_1, p_3, p_1, p'_1)$.

Decoding: Use $T^n_\epsilon(P_Z)$ to denote the set of jointly strong typical n-sequence with respect to the p.d.f. $p(z)$. Without loss of generality, assume that $(w_0, w_1, p_3, p_1) = (1, 1, 1, 1)$ is sent and w'_1 and p'_1 can be arbitrary. The receivers decode as follows:

1) Receiver 1 uses joint typical decoding of $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{x}, \mathbf{y}_1\}$ to find the indices (w_0, w_1, p_3, p_1).
2) Receiver 2 uses indirect decoding of $\mathbf{u}_2$ [10] to find the index w_0. Once this is known, $\mathbf{u}_1$ is also found. Then, receiver 2 uses joint typical decoding of $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{y}_2\}$ to find w_1.

3) Receiver 3 uses indirect decoding of $\mathbf{u}_3$ to find the index w_0.

At receiver 1, the decoder seeks the indices (w_0, w_1, p_3, p_1) so that

$$(\mathbf{u}_1(w_0), \mathbf{u}_2(w_0, w_1, w_1'), \mathbf{u}_3(p_3), \mathbf{x}(w_0, w_1, w_1', p_3, p_1, p_1'), \mathbf{y}_1) \in T_\epsilon^n(P_{U_1U_2U_3XY_1}). \tag{19}$$

If there is none or more than one possible codeword, an error is declared. The possible error events are as follows:

a) $\mathtt{E}_1 : (w_0, w_1, w_1', p_3, p_1, p_1') = (1, 1, w_1', 1, 1, p_1')$ but $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{x}$ are not jointly typical with $\mathbf{y}$. By the properties of strong typical sequences [17], $\Pr\{\mathtt{E}_1\} \leq \epsilon'$, where $\epsilon' \to 0$ for large n.

b) $\mathtt{E}_2 : w_0 \neq 1$ and arbitrary w_1, p_3, p_1, with $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{x}$ jointly typical with $\mathbf{y}_1$. Then, we have

$$\begin{aligned}\Pr\{\mathtt{E}_2\} \leq & \sum_{\substack{w_0 \neq 1 \\ w_1, p_3, p_1, w_1', p_1'}} \Pr\left\{(\mathbf{U}_1(w_0), \mathbf{U}_2(w_0, w_1, w_1'), \mathbf{U}_3(p_3), \mathbf{X}(w_0, w_1, w_1', p_3, p_1, p_1'), \mathbf{y}_1) \in T_\epsilon^n(P_{U_1U_2U_3XY_1})\right\} \\ & \leq 2^{n(R_0 + R_{1e} + R_1' + P_{1e} + P_1' + P_3)} 2^{-n(I(U_1, U_2, U_3, X; Y_1) - 2\delta)},\end{aligned} \tag{20}$$

where $\delta \to 0$ as $\epsilon \to 0$ for n sufficiently large. For $\Pr\{\mathtt{E}_2\} \leq \epsilon'$, we require

$$R_0 + R_{1e} + R_1' + P_{1e} + P_1' + P_3 < I(U_1, U_2, U_3, X; Y_1) = I(X; Y_1) \tag{21}$$

since $I(U_1, U_2, U_3; Y_1|X) = 0$ by the Markov chain condition

$$U_1 \to (U_2, U_3) \to X \to Y_1. \tag{22}$$

c) $\mathtt{E}_3 : w_0 = 1, w_1 \neq 1$ and arbitrary p_3, p_1, with $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{x}$ jointly typical with $\mathbf{y}_1$. Then, we have

$$\begin{aligned}\Pr\{\mathtt{E}_3\} \leq & \sum_{\substack{w_1 \neq 1 \\ p_3, p_1, w_1', p_1'}} \Pr\left\{(\mathbf{u}_1(1), \mathbf{U}_2(1, w_1, w_1'), \mathbf{U}_3(p_3), \mathbf{X}(1, w_1, w_1', p_3, p_1, p_1'), \mathbf{y}_1) \in T_\epsilon^n(P_{U_1U_2U_3XY_1})\right\} \\ & \leq 2^{n(R_{1e} + R_1' + P_{1e} + P_1' + P_3)} 2^{-n(I(U_2, U_3, X; Y_1|U_1) - 2\delta)}.\end{aligned} \tag{23}$$

For $\Pr\{\mathtt{E}_3\} \leq \epsilon'$, we require

$$\begin{aligned}R_{1e} + R_1' + P_{1e} + P_1' + P_3 &< I(U_2, U_3, X; Y_1|U_1) \\ &= I(X; Y_1|U_1) + I(U_2, U_3; Y_1|X, U_1) = I(X; Y_1|U_1),\end{aligned} \tag{24}$$

where the second line is due to $U_1 \to (U_2, U_3) \to X \to Y_1$.

d) $\mathtt{E}_4 : w_0 = 1, w_1 = 1, p_3 \neq 1$ and arbitrary p_1, with $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{x}$ jointly typical with $\mathbf{y}_1$. Then, we have

$$\begin{aligned}\Pr\{\mathtt{E}_4\} \leq & \sum_{\substack{p_3 \neq 1 \\ p_1, w_1', p_1'}} \Pr\left\{(\mathbf{u}_1(1), \mathbf{u}_2(1, 1, w_1'), \mathbf{U}_3(p_3), \mathbf{X}(1, 1, w_1', p_3, p_1, p_1'), \mathbf{y}_1) \in T_\epsilon^n(P_{U_1U_2U_3XY_1})\right\} \\ & \leq 2^{n(P_{1e} + P_1' + P_3)} 2^{-n(I(U_3, X; Y_1|U_1, U_2) - 2\delta)}.\end{aligned} \tag{25}$$

For $\Pr\{\mathtt{E}_4\} \leq \epsilon'$, we require

$$\begin{aligned} P_{1e} + P_1' + P_3 < I(U_3, X; Y_1|U_1, U_2) &= I(X; Y_1|U_1, U_2) + I(U_3; Y_1|U_1, U_2, X) \\ &\overset{(a)}{=} I(X; Y_1|U_2) + I(U_3; Y_1|U_2, X) \\ &\overset{(b)}{=} I(X; Y_1|U_2), \end{aligned} \tag{26}$$

where the first term in (a) is due to $U_1 \to U_2 \to X \to Y_1$ and the second term is due to $U_1 \to (U_2, U_3) \to X \to Y_1$, and (b) is due to $U_3 \to (U_2, X) \to Y_1$.

e) $\mathtt{E}_5 : w_0 = 1, w_1 = 1, p_3 = 1, p_1 \neq 1$ with $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{x}$ jointly typical with $\mathbf{y}_1$. Then, we have

$$\begin{aligned} \Pr\{\mathtt{E}_5\} &\leq \sum_{\substack{p_1 \neq 1 \\ p_1'}} \Pr\left\{(\mathbf{u}_1(1), \mathbf{u}_2(1, 1, w_1'), \mathbf{u}_3(1), \mathbf{X}(1, 1, w_1', 1, p_1, p_1'), \mathbf{y}_1) \in T_\epsilon^n(P_{U_1U_2U_3XY_1})\right\} \\ &\leq 2^{n(P_{1e}+P_1')} 2^{-n(I(X;Y_1|U_1,U_2,U_3)-2\delta)}. \end{aligned} \tag{27}$$

For $\Pr\{\mathtt{E}_5\} \leq \epsilon'$, we require

$$P_{1e} + P_1' < I(X; Y_1|U_1, U_2, U_3) = I(X; Y_1|U_2, U_3) \tag{28}$$

where the equality is due to $U_1 \to (U_2, U_3) \to X \to Y_1$.

e) $\mathtt{E}_6 : w_0 = 1, w_1 \neq 1, p_3 = 1$ and p_1 arbitrary with $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{x}$ jointly typical with $\mathbf{y}_1$. Then, we have

$$\begin{aligned} \Pr\{\mathtt{E}_6\} &\leq \sum_{\substack{w_1 \neq 1 \\ p_1, w_1', p_1'}} \Pr\left\{(\mathbf{u}_1(1), \mathbf{U}_2(1, w_1, w_1'), \mathbf{u}_3(1), \mathbf{X}(1, w_1, w_1', 1, p_1, p_1'), \mathbf{y}_1) \in T_\epsilon^n(P_{U_1U_2U_3XY_1})\right\} \\ &\leq 2^{n(R_{1e}+R_1'+P_{1e}+P_1')} 2^{-n(I(U_2,X;Y_1|U_1,U_3)-2\delta)}. \end{aligned} \tag{29}$$

For $\Pr\{\mathtt{E}_6\} \leq \epsilon'$, we require

$$\begin{aligned} R_{1e} + R_1' + P_{1e} + P_1' < I(U_2, X; Y_1|U_1, U_3) &= I(X; Y_1|U_1, U_3) + I(U_2; Y_1|U_1, U_3, X) \\ &\overset{(a)}{=} I(X; Y_1|U_3) + I(U_2; Y_1|U_3, X) \\ &\overset{(b)}{=} I(X; Y_1|U_3), \end{aligned} \tag{30}$$

where the first term of (a) is due to $U_1 \to U_3 \to X \to Y_1$ and the second term of (a) and (b) are due to $U_1 \to U_2 \to (U_3, X) \to Y_1$. Consequently, under the conditions (21), (24), (26), (28), (30) listed above, the error probability at receiver 1 is less than $\sum_{i=1}^{6} \Pr\{\mathtt{E}_i\} \leq 6\epsilon'$.

Now, assume that $(w_0, q_2) = (1, 1)$ is sent to receiver 2. At receiver 2, the decoder first finds w_0 by indirect decoding, then finds w_1 by joint typical decoding. The error events at receiver 2 may be divided into:

a) $\mathtt{E}_1' : (w_0, q_2) = (1, 1)$ but $\mathbf{u}_2$ is not jointly typical with $\mathbf{y}_2$ (indirect decoding). In this case, by the properties of strong typical sequences, we have $\Pr\{\mathtt{E}_1'\} \leq \epsilon'$.

b) $\mathtt{E}_2'$: $w_0 \neq 1$, q_2 arbitrary and $\mathbf{u}_2$ is jointly typical with $\mathbf{y}_2$ (indirect decoding). This is the same as receiver 2 trying to estimate w_0 such that $(\mathbf{u}_2(w_0, q_2), \mathbf{y}_3) \in T_\epsilon^n(P_{U_2Y_2})$ for any $q_2 \in \{1, \ldots, 2^{nQ_2}\}$. We have

$$\Pr\{\mathtt{E}_2'\} \leq \sum_{w_0 \neq 1} \sum_{q_2} \Pr\{(\mathbf{U}_2(w_0, q_2), \mathbf{y}_2) \in T_\epsilon^n(P_{U_2Y_2})\} \leq 2^{n(R_0+Q_2)} 2^{-n(I(U_2;Y_2)-2\delta)}. \tag{31}$$

Then, for $\Pr\{\mathtt{E}_2'\} \leq \epsilon'$, we need

$$R_0 + Q_2 < I(U_2; Y_2). \tag{32}$$

c) $\mathtt{E}_3'$: $w_0 = 1$, $q_2 \neq 1$, and $\mathbf{u}_1$, $\mathbf{u}_2$ are jointly typical with $\mathbf{y}_2$. Then, we have

$$\Pr\{\mathtt{E}_3'\} \leq \sum_{q_2} \Pr\{(\mathbf{u}_1(1), \mathbf{U}_2(1, q_2), \mathbf{y}_2) \in T_\epsilon^n(P_{U_1U_2Y_2})\} \leq 2^{nQ_2} 2^{-n(I(U_2;Y_2|U_1)-2\delta)}. \tag{33}$$

Then, for $\Pr\{\mathtt{E}_2'\} \leq \epsilon'$, we need

$$Q_2 < I(U_2; Y_2|U_1). \tag{34}$$

Thus, under the conditions (32) and (34), the error probability at receiver 2 is less than $\sum_{i=1}^{3} \Pr\{\mathtt{E}_i'\} \leq 3\epsilon'$.

At receiver 3, indirect decoding is used, so that the decoder estimates w_0 such that $(\mathbf{u}_3(w_0, q_3), \mathbf{y}_3) \in T_\epsilon^n(P_{U_3Y_3})$ for any $q_3 \in \{1, \ldots, 2^{nQ_3}\}$. Assuming that $(w_0, q_3) = (1, 1)$ is sent, we require

$$R_0 + Q_3 < I(U_3; Y_3), \tag{35}$$

for the error probability at receiver 3 to be small for n sufficiently large.

In addition to the decoding conditions above, we require that

$$P_{1e} + P_1' > I(X; Y_1|U_2), \tag{36}$$

which is a consequence of setting $P_1' = I(X; Y_1|U_2) - \delta_1$ as the partition size.

Combining (18), (21), (24), (26), (28), (30), (32), (34), (35) and (36) using Fourier-Motzkin elimination with $R_1 = R_{1e} + R_1'$, $R_2 = R_{2e} + R_2'$, $R_{2e} = P_{1e} + P_3$, we can obtain the inner bound to the secrecy capacity region in Theorem 1 as well as condition (8), which completes the proof.

B. Proof of Achievability for 3-Receiver BC with 2 Degraded Message Sets (Type 1)

Here, we outline the proof of achievability for the Type 1 3-receiver BC with 2 degraded message sets and secrecy constraints. The coding scheme largely follows that for the 3 degraded message sets case, but with the key difference being the assignment of the message W_1 using the auxiliary codewords. Specifically, instead of encoding W_1 using the auxiliary codeword $\mathbf{U}_2$ and W_2 using $\mathbf{U}_3$ and $\mathbf{X}$ as in the 3 degraded message sets case, here, W_1 is encoded using $\mathbf{U}_2$, $\mathbf{U}_3$ and $\mathbf{X}$. We can use the same code partitions and sizes of the partitions for security as in the 3 degraded message sets case, even for this different coding scheme.

Codebook generation: Let us define $R_1 \triangleq R_{1e} + R_1'$, $R_{1e} \triangleq P_{1e} + P_{2e} + P_3$, $R_1' \triangleq P_1' + P_2'$, and

$$\begin{cases} P_1' \triangleq I(X;Y_3|U_2) - \delta_1, \\ P_2' \triangleq I(U_2;Y_3|U_1) - \delta_1, \end{cases} \tag{37}$$

where $\delta_1 > 0$ and is small for n sufficiently large.

The code generation follows the same way as in Section IV-A, except that we randomly partition the sequences, $\mathbf{U}_2(w_0, q_2)$, into $2^{n\tilde{P}_2}$ equally-sized bins, and $\mathbf{U}_3(w_0, q_3)$, into $2^{n\tilde{P}_3}$ equally-sized bins, where $\tilde{P}_2 = P_{2e} + P_2' + P_2^\dagger$ and $\tilde{P}_3 = P_3 + P_3^\dagger$. The $\mathbf{U}_2$ codewords undergo a double partition while $\mathbf{U}_3$ undergo a single partition. Then, for each product bin (p_2, p_3) contains the joint typical pair $(\mathbf{u}_2(p_2, p_2', p_2^\dagger), \mathbf{u}_3(p_3, p_3^\dagger))$ for $p_2 \in \{1, \ldots, 2^{nP_{2e}}\}$, $p_2' \in \{1, \ldots, 2^{P_2'}\}$, $p_2^\dagger \in \{1, \ldots, 2^{nP_2^\dagger}\}$, $p_3 \in \{1, \ldots, 2^{nP_3}\}$, $p_3^\dagger \in \{1, \ldots, 2^{nP_3^\dagger}\}$ with high probability

$$\begin{aligned} P_{2e} + P_2' &\leq Q_2, \\ P_3 &\leq Q_3, \\ P_{2e} + P_2' + P_3 &\leq Q_2 + Q_3 - I(U_2;U_3|U_1). \end{aligned} \tag{38}$$

As before, for each joint typical pair $(\mathbf{u}_2(p_2, p_2'), \mathbf{u}_3(p_3))$ corresponding to the product bin (p_2, p_2', p_3), generate $2^{n\tilde{P}_1}$ sequences of codewords $\mathbf{X}(w_0, p_1, p_1', p_2, p_2', p_3)$, where $\tilde{P}_1 = P_{1e} + P_1'$, for $p_1 \in \{1, \ldots, 2^{nP_{1e}}\}$ and $p_1' \in \{1, \ldots, 2^{nP_1'}\}$, uniformly and randomly over the set of conditionally typical $\mathbf{X}$ sequences. The $2^{n\tilde{P}_1}$ codewords are partitioned into $2^{nP_{1e}}$ subcodes with $2^{nP_1'}$ codewords within the subcodes.

Encoding: To send (w_0, w_1), express w_1 by (p_1, p_2, p_3) and send the codeword $\mathbf{x}(w_0, p_1, p_1', p_2, p_2', p_3)$.

Decoding: Assume that $(w_0, p_1, p_2, p_3) = (1, 1, 1, 1)$ is sent and p_1', p_2' can be arbitrary. The receivers decode the messages as follows:

1) Receiver 1 uses joint typical decoding of $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{x}, \mathbf{y}_1\}$ to find the indices (w_0, p_1, p_2, p_3).
2) Receiver 2 uses indirect decoding of $\mathbf{u}_2$ to find the index w_0.
3) Receiver 3 uses indirect decoding of $\mathbf{u}_3$ to find the index w_0.

At receiver 1, the decoder seeks the message so that

$$(\mathbf{u}_1(w_0), \mathbf{u}_2(p_2, p_2'), \mathbf{u}_3(p_3), \mathbf{x}(w_0, p_1, p_1', p_2, p_2', p_3), \mathbf{y}_1) \in T_\epsilon^n(P_{U_1 U_2 U_3 X Y_1}). \tag{39}$$

The error events at receiver 1 can be classified into:

a) $\mathtt{E}_1 : (w_0, p_1, p_1', p_2, p_2', p_3) = (1, 1, p_1', 1, p_2', 1)$ but $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{x}$ are not jointly typical with $\mathbf{y}_1$. In this case, we have $\Pr\{\mathtt{E}_1\} \leq \epsilon \to 0$ for large n.

b) $\mathtt{E}_2 : w_0 \neq 1$, with arbitrary p_1, p_2 and p_3, but $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ and $\mathbf{x}$ are jointly typical with $\mathbf{y}_1$. For $\Pr\{\mathtt{E}_2\} \leq \epsilon \to 0$ with n sufficiently large to be true, we then need

$$R_0 + P_{1e} + P_1' + P_{2e} + P_2' + P_3 < I(X;Y_1). \tag{40}$$

c) $\mathtt{E}_3$: $w_0 = 1$, $p_2, p_3 \neq 1$, and p_1 arbitrary, but $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{x}$ are jointly typical with $\mathbf{y}_1$. For $\Pr\{\mathtt{E}_3\} \leq \epsilon \to 0$ with n sufficiently large to be true, we require

$$P_{1e} + P_1' + P_{2e} + P_2' + P_3 < I(X; Y_1|U_1). \tag{41}$$

d) $\mathtt{E}_4$: $w_0 = 1$, $p_2 = 1$, $p_3 \neq 1$, and p_1 arbitrary, but $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{x}$ are jointly typical with $\mathbf{y}_1$. Then, for $\Pr\{\mathtt{E}_4\} \leq \epsilon \to 0$ with n sufficiently large to be true, we need

$$P_{1e} + P_1' + P_3 < I(X; Y_1|U_2). \tag{42}$$

e) $\mathtt{E}_5$: $w_0 = 1$, $p_2 \neq 1$, $p_3 = 1$, and p_1 arbitrary, but $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{x}$ are jointly typical with $\mathbf{y}_1$. Then, for $\Pr\{\mathtt{E}_5\} \leq \epsilon \to 0$ with n sufficiently large to be true, we need

$$P_{1e} + P_1' + P_{2e} + P_2' < I(X; Y_1|U_3). \tag{43}$$

f) $\mathtt{E}_6$: $w_0 = 1$, $p_2 = 1$, $p_3 = 1$ and $p_1 \neq 1$, but $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{x}$ are jointly typical with $\mathbf{y}_1$. Then, for $\Pr\{\mathtt{E}_6\} \leq \epsilon \to 0$ with n sufficiently large to be true, we require

$$P_{1e} + P_1' < I(X; Y_1|U_2, U_3). \tag{44}$$

The error probability at receiver 1 is therefore less than $\sum_{i=1}^{6} \Pr\{\mathtt{E}_i\} \leq 6\epsilon$.

At receivers 2 and 3, assuming that $(w_0, q_2) = (w_0, q_3) = (1, 1)$ is sent, we require

$$\begin{cases} R_0 + Q_2 < I(U_2; Y_2), \\ R_0 + Q_3 < I(U_3; Y_3), \end{cases} \tag{45}$$

for the error probabilities tending to 0 for n sufficiently large. We additionally have

$$P_{1e} + P_1' > I(X; Y_1|U_2), \tag{46}$$

which is a consequence of setting $P_1' = I(X; Y_1|U_2) - \delta_1$ as the partition size.

Combining (38) and (40) to (45) and (46) by using Fourier-Motzkin elimination with $R_1 = R_{1e} + R_1'$, $R_{1e} = P_{1e} + P_{2e} + P_3$, we can obtain the rate region in Theorem 2 and the conditions (13).

C. Equivocation Calculation for 3-Receiver BC with 3 Degraded Message Sets

In this section, we show that the equivocation rate for the 3-receiver BC with 3 degraded message sets satisfies the security conditions in (5). That is, we shall derive the bounds for $H(W_1|\mathbf{Y}_3)$, $H(W_2|\mathbf{Y}_3)$ and $H(W_1, W_2|\mathbf{Y}_3)$. In the analysis, we shall make use of the following relation very frequently

$$H(U, V) = H(U) + H(V|U). \tag{47}$$

For the message W_1, the equivocation can be bounded in two ways, which respectively correspond to whether $\mathbf{U}_2$ is the codeword sent to Y_2 or $\mathbf{X}$ is the codeword sent to Y_1. For the former case, we have

$$\begin{aligned}
H(W_1|\mathbf{Y}_3) &\geq H(W_1|\mathbf{Y}_3, \mathbf{U}_1)\\
&\overset{(a)}{=} H(W_1, \mathbf{Y}_3|\mathbf{U}_1) - H(\mathbf{Y}_3|\mathbf{U}_1)\\
&\overset{(b)}{=} H(W_1, \mathbf{U}_2, \mathbf{Y}_3|\mathbf{U}_1) - H(\mathbf{U}_2|W_1, \mathbf{U}_1, \mathbf{Y}_3) - H(\mathbf{Y}_3|\mathbf{U}_1)\\
&\geq H(\mathbf{U}_2|\mathbf{U}_1) + [H(\mathbf{Y}_3|\mathbf{U}_2, \mathbf{U}_1) - H(\mathbf{Y}_3|\mathbf{U}_1)] - H(\mathbf{U}_2|W_1, \mathbf{U}_1, \mathbf{Y}_3)\\
&= H(\mathbf{U}_2|\mathbf{U}_1) - I(\mathbf{U}_2; \mathbf{Y}_3|\mathbf{U}_1) - H(\mathbf{U}_2|W_1, \mathbf{U}_1, \mathbf{Y}_3), \qquad (48)
\end{aligned}$$

where (a) is by (47), and (b) has first two terms by (47). Now, we can bound each term in (48) separately. For the first term, given $\mathbf{u}_1$, $\mathbf{U}_2$ has $2^{nI(U_2;Y_2|U_1)}$ codewords with equal probability. As such,

$$H(\mathbf{U}_2|\mathbf{U}_1) = nI(U_2; Y_2|U_1) - n\delta_1', \qquad (49)$$

where $\delta_1' > 0$ and is small for n sufficiently large. The second term can be bounded by [3]

$$I(\mathbf{U}_2; \mathbf{Y}_3|\mathbf{U}_1) \leq nI(U_2; Y_3|U_1) + n\delta', \qquad (50)$$

where $\delta' > 0$ and is small for n sufficiently large. For the third term, by Fano's inequality, we have

$$\frac{1}{n}H(\mathbf{U}_2|W_1, \mathbf{U}_1, \mathbf{Y}_3) \leq \frac{1}{n}(1 + \lambda(w_1') \log R_1') \triangleq \epsilon_{2,n}', \qquad (51)$$

where $\epsilon_{2,n}' \to 0$ for n sufficiently large.

To show that $\lambda(w_1') \leq 2\kappa$ where $\kappa \to 0$ for n sufficiently large so that (51) holds, consider decoding at the wiretapper and the codebook with rate R_1' to be decoded at the wiretapper with error probability $\lambda(w_1')$. Let $W_1 = w_1$ and $W_0 = w_0$ be fixed. We note that the wiretapper decodes $\mathbf{U}_2$ first as it will then use this knowledge to decode $\mathbf{X}$ later. The wiretapper decodes $\mathbf{U}_2$ given $W_1 = w_1$ and $\mathbf{U}_1 = \mathbf{u}_1$, by finding the index w_1', so that

$$(\mathbf{u}_1(w_0), \mathbf{u}_2(w_0, w_1, w_1'), \mathbf{y}_3) \in T_\epsilon^n(P_{U_1U_2Y_3}). \qquad (52)$$

If there is none or more than one possible codeword, an error is declared. Now, define the event

$$\mathtt{E}_1^{(Y_3)}(w_1') \triangleq \{\mathbf{u}_1(w_0), \mathbf{U}_2(w_0, w_1, w_1'), \mathbf{y}_3 \in T_\epsilon^n(P_{U_1U_2Y_3})\}. \qquad (53)$$

Then, assuming that $\mathbf{u}_2(w_0, w_1, 1)$ is sent,

$$\lambda(w_1') \leq \Pr\left\{\left(\mathtt{E}_1^{(Y_3)}(1)\right)^c\right\} + \sum_{w_1'} \Pr\left\{\mathtt{E}_1^{(Y_3)}(1)\right\} \leq \kappa + 2^{nR_1'}2^{-n(I(U_2;Y_3|U_1)-2\delta)}, \qquad (54)$$

where $\delta \to 0$ as $\epsilon \to 0$ for n sufficiently large. Thus, since we have chosen $R_1' = I(U_2; Y_3|U_1) - \delta_1$ for the double-binning partition, we get $\lambda(w_1') \leq 2\kappa$ for $\delta_1 > 2\delta$ and (51) holds. Substituting (49)–(51) into (48), we have $H(W_1|\mathbf{Y}_3) \geq nR_{1e} - n\epsilon_1$, where $\epsilon_1 = \delta'_1 + \delta' + \epsilon'_{2,n}$, and hence the equivocation rate satisfies the first condition in (5a).

For message W_1 sent using $\mathbf{X}$ to Y_1, we have

$$\begin{aligned} H(W_1|\mathbf{Y}_3) &\geq H(W_1, \mathbf{Y}_3|\mathbf{U}_1) - H(\mathbf{Y}_3|\mathbf{U}_1) \\ &= H(W_1, \mathbf{X}, \mathbf{Y}_3|\mathbf{U}_1) - H(\mathbf{X}|W_1, \mathbf{U}_1, \mathbf{Y}_3) - H(\mathbf{Y}_3|\mathbf{U}_1) \\ &\geq H(\mathbf{X}|\mathbf{U}_1) + H(\mathbf{Y}_3|\mathbf{U}_1, \mathbf{X}) - H(\mathbf{Y}_3|\mathbf{U}_1) - H(\mathbf{X}|W_1, \mathbf{U}_1, \mathbf{Y}_3) \\ &\geq H(\mathbf{X}|\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3) + H(\mathbf{Y}_3|\mathbf{U}_1, \mathbf{U}_2, \mathbf{X}) - H(\mathbf{Y}_3|\mathbf{U}_1) - H(\mathbf{U}_2, \mathbf{X}|W_1, \mathbf{U}_1, \mathbf{Y}_3) \\ &= H(\mathbf{X}|\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3) - I(\mathbf{U}_2, \mathbf{X}; \mathbf{Y}_3|\mathbf{U}_1) - H(\mathbf{U}_2|W_1, \mathbf{U}_1, \mathbf{Y}_3) - H(\mathbf{X}|W_1, \mathbf{U}_1, \mathbf{U}_2, \mathbf{Y}_3). \quad (55) \end{aligned}$$

For the first term in (55), given $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$, $\mathbf{X}$ has $2^{nI(X;Y_1|U_2,U_3,U_1)}$ codewords with equal probability. Then,

$$\begin{aligned} H(\mathbf{X}|\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3) &= nI(X; Y_1|U_2, U_3, U_1) - n\delta'_1 \\ &= nI(X; Y_1|U_2) - n\delta'_1 \text{ or } nI(X; Y_1|U_3) - n\delta'_1. \quad (56) \end{aligned}$$

The last equalities are due to $I(X; Y_1|U_2, U_3, U_1) = I(X; Y_1|U_1) - I(U_2, U_3; Y_1|U_1)$ and

$$I(U_2, U_3; Y_1|U_1) = I(U_2; Y_1|U_1) + I(U_3; Y_1|U_2, U_1) = I(U_2; Y_1|U_1), \quad (57)$$

$$I(U_2, U_3; Y_1|U_1) = I(U_3; Y_1|U_1) + I(U_2; Y_1|U_3, U_1) = I(U_3; Y_1|U_1), \quad (58)$$

where the above equalities are due to the Markov chain conditions (9). Thus, for this case, we choose

$$H(\mathbf{X}|\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3) = nI(X; Y_1|U_3) - n\delta'_1. \quad (59)$$

The second term in (55) can be bounded as

$$I(\mathbf{U}_2, \mathbf{X}; \mathbf{Y}_3|\mathbf{U}_1) = I(\mathbf{U}_2; \mathbf{Y}_3|\mathbf{U}_1) + I(\mathbf{X}; \mathbf{Y}_3|\mathbf{U}_2, \mathbf{U}_1) \leq nI(U_2; Y_3|U_1) + nI(X; Y_3|U_2) + 2n\delta'. \quad (60)$$

The third term in (55) may be bounded using Fano's inequality as in (51). The fourth term can also be bounded using Fano's inequality, by which we have

$$\frac{1}{n}H(\mathbf{X}|W_1, \mathbf{U}_1, \mathbf{U}_2, \mathbf{Y}_3) \leq \frac{1}{n}(1 + \lambda(p'_1) \log P'_1) \triangleq \epsilon'_{1,n}, \quad (61)$$

where $\epsilon'_{1,n} \to 0$ for n sufficiently large. To show that $\lambda(p'_1) \leq 2\kappa$ so that (61) holds, assume that wiretapper Y_3 knows $\mathbf{U}_2 = \mathbf{u}_2$, $\mathbf{U}_1 = \mathbf{u}_1$ and decodes $\mathbf{x}(w_0, w_1, w'_1, p_3, p_1, p'_1)$ by finding the index p'_1, so that

$$(\mathbf{u}_1(w_0), \mathbf{u}_2(w_0, w_1, w'_1), \mathbf{u}_3(p_3), \mathbf{x}(w_0, w_1, w'_1, p_3, p_1, p'_1), \mathbf{y}_3) \in T^n_\epsilon(P_{U_1U_2U_3XY_3}). \quad (62)$$

If there is none or more than one possible codeword, an error is declared. Define the event

$$\mathtt{E}_2^{(Y_3)}(p'_1) \triangleq \{\mathbf{u}_1(w_0), \mathbf{u}_2(w_0, w_1, w'_1), \mathbf{U}_3(p_3), \mathbf{X}(w_0, w_1, w'_1, p_3, p_1, p'_1), \mathbf{y}_3 \in T^n_\epsilon(P_{U_1U_2U_3XY_3})\}, \quad (63)$$

where w_0, w_1, w_1' are known. Assuming that $\mathbf{x}(w_0, w_1, w_1', p_3, p_1, 1)$ is sent, we then have

$$\lambda(p_1') \leq \Pr\left\{\left(\mathrm{E}_2^{(Y_3)}(1)\right)^c\right\} + \sum_{p_1'} \Pr\left\{\mathrm{E}_2^{(Y_3)}(1)\right\} \leq \kappa + 2^{nP_1'} 2^{-n(I(X;Y_3|U_1,U_2)-2\delta)}, \tag{64}$$

where $\delta \to 0$ as $\epsilon \to 0$ for n sufficiently large. Since we have chosen $P_1' = I(X;Y_3|U_2) - \delta_1$, we obtain $\lambda(p_1') \leq 2\kappa$ for $\delta_1 > 2\delta$. Thus, (61) holds and substituting (59), (60), (51), (61) into (55), we have $H(W_1|\mathbf{Y}_3) \geq nR_{1e} - n\tilde{\epsilon}_1$, where $n\tilde{\epsilon}_1 = \delta_1' + 2\delta' + \epsilon_{1,n}' + \epsilon_{2,n}'$ is small for n sufficiently large, so the second condition in (5a) is satisfied.

For the message W_2, the equivocation can be bounded by

$$\begin{aligned} H(W_2|\mathbf{Y}_3) &\geq H(W_2|\mathbf{Y}_3, \mathbf{U}_1, \mathbf{U}_2) \\ &= H(W_2, \mathbf{Y}_3|\mathbf{U}_1, \mathbf{U}_2) - H(\mathbf{Y}_3|\mathbf{U}_1, \mathbf{U}_2) \\ &= H(W_2, \mathbf{X}, \mathbf{Y}_3|\mathbf{U}_1, \mathbf{U}_2) - H(\mathbf{X}|W_2, \mathbf{U}_1, \mathbf{U}_2, \mathbf{Y}_3) - H(\mathbf{Y}_3|\mathbf{U}_1, \mathbf{U}_2) \\ &\geq H(\mathbf{X}|\mathbf{U}_1, \mathbf{U}_2) + [H(\mathbf{Y}_3|\mathbf{U}_1, \mathbf{U}_2, \mathbf{X}) - H(\mathbf{Y}_3|\mathbf{U}_1, \mathbf{U}_2)] - H(\mathbf{X}|W_2, \mathbf{U}_1, \mathbf{U}_2, \mathbf{Y}_3) \\ &\geq H(\mathbf{X}|\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3) - I(\mathbf{X}; \mathbf{Y}_3|\mathbf{U}_1, \mathbf{U}_2) - H(\mathbf{X}|W_2, \mathbf{U}_1, \mathbf{U}_2, \mathbf{Y}_3). \end{aligned} \tag{65}$$

For the first term in (65), given $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{X}$ has $2^{nI(X;Y_1|U_2,U_3,U_1)}$ codewords with equal probability. Thus,

$$H(\mathbf{X}|\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3) = nI(X;Y_1|U_2) - n\delta_1', \tag{66}$$

as discussed in the obtaining of (56). The second term is bounded by

$$I(\mathbf{X}; \mathbf{Y}_3|\mathbf{U}_2, \mathbf{U}_3) \leq nI(X;Y_3|U_2, U_3) + n\delta', \tag{67}$$

where $\delta' > 0$ and is small for n sufficiently large. For the third term, by Fano's inequality, we have

$$\frac{1}{n} H(\mathbf{X}|W_2, \mathbf{U}_2, \mathbf{U}_3, \mathbf{Y}_3) \leq \frac{1}{n}(1 + \lambda(p_1') \log P_1') \triangleq \epsilon_{3,n}', \tag{68}$$

where $\epsilon_{3,n}' \to 0$ for n sufficiently large. To show that $\lambda(p_1') \leq 2\kappa$ so that (68) holds, since the wiretapper knows W_2, we can assume that wiretapper Y_3 knows $\mathbf{U}_3 = \mathbf{u}_3$, $\mathbf{U}_2 = \mathbf{u}_2$, $\mathbf{U}_1 = \mathbf{u}_1$ and decodes $\mathbf{x}(w_0, w_1, w_1', p_3, p_1, p_1')$ by finding the index p_1', such that

$$(\mathbf{u}_1(w_0), \mathbf{u}_2(w_0, w_1, w_1'), \mathbf{u}_3(p_3), \mathbf{x}(w_0, w_1, w_1', p_3, p_1, p_1'), \mathbf{y}_3) \in T_\epsilon^n(P_{U_1U_2U_3XY_3}). \tag{69}$$

If there is none or more than one possible codeword, an error is declared. Define the event

$$\mathrm{E}_2^{(Y_3)}(p_1') \triangleq \{\mathbf{u}_1(w_0), \mathbf{u}_2(w_0, w_1, w_1'), \mathbf{u}_3(p_3), \mathbf{X}(w_0, w_1, w_1', p_3, p_1, p_1'), \mathbf{y}_3 \in T_\epsilon^n(P_{U_1U_2U_3XY_3})\}, \tag{70}$$

where w_0, w_1, w_1', p_3 are known. Assuming that $\mathbf{x}(w_0, w_1, w_1', p_3, p_1, 1)$ is sent, we then have

$$\lambda(p_1') \leq \Pr\left\{\left(\mathrm{E}_2^{(Y_3)}(1)\right)^c\right\} + \sum_{p_1'} \Pr\left\{\mathrm{E}_2^{(Y_3)}(1)\right\} \leq \kappa + 2^{nP_1'} 2^{-n(I(X;Y_3|U_1,U_2,U_3)-2\delta)}, \tag{71}$$

where $\delta \to 0$ as $\epsilon \to 0$ for n sufficiently large. Since we have chosen

$$P_1' = I(X;Y_3|U_2) - \delta_1 = I(X;Y_3|U_1,U_2,U_3) - \delta_1, \tag{72}$$

we obtain $\lambda(p_1') \leq 2\kappa$ for $\delta_1 > 2\delta$ and (68) holds. Substituting (66), (67) and (68) into (65), we have $H(W_2|\mathbf{Y}_3) \geq nR_{2e} - n\epsilon_2$, where $\epsilon_2 = \delta_1' + \delta' + \epsilon_{3,n}'$, and the equivocation rate satisfies (5b).

For the combined message (W_1, W_2), we have

$$\begin{aligned} H(W_1,W_2|\mathbf{Y}_3) &\geq H(W_1,W_2|\mathbf{Y}_3,\mathbf{U}_1) \\ &= H(W_1,W_2,\mathbf{Y}_3|\mathbf{U}_1) - H(\mathbf{Y}_3|\mathbf{U}_1) \\ &= H(W_1,W_2,\mathbf{X},\mathbf{Y}_3|\mathbf{U}_1) - H(\mathbf{X}|W_1,W_2,\mathbf{U}_1,\mathbf{Y}_3) - H(\mathbf{Y}_3|\mathbf{U}_1) \\ &\geq H(\mathbf{X}|\mathbf{U}_1) + H(\mathbf{Y}_3|\mathbf{U}_1,\mathbf{X}) - H(\mathbf{Y}_3|\mathbf{U}_1) - H(\mathbf{U}_2,\mathbf{X}|W_1,W_2,\mathbf{U}_1,\mathbf{Y}_3) \\ &\geq H(\mathbf{X}|\mathbf{U}_1) + [H(\mathbf{Y}_3|\mathbf{U}_1,\mathbf{U}_2,\mathbf{X}) - H(\mathbf{Y}_3|\mathbf{U}_1)] - H(\mathbf{U}_2|W_1,W_2,\mathbf{U}_1,\mathbf{Y}_3) \\ &\quad - H(\mathbf{X}|W_1,W_2,\mathbf{U}_1,\mathbf{U}_2,\mathbf{Y}_3). \end{aligned} \tag{73}$$

For the first term, we have

$$H(\mathbf{X}|\mathbf{U}_1) = nI(X;Y_1|U_1) - n\delta_1'. \tag{74}$$

The second term can be bounded by

$$I(\mathbf{U}_2,\mathbf{X};\mathbf{Y}_3|\mathbf{U}_1) \leq nI(U_2;Y_3|U_1) + nI(X;Y_3|U_2) + 2n\delta'. \tag{75}$$

The fourth and fifth terms are, respectively,

$$\frac{1}{n}H(\mathbf{U}_2|W_1,W_2,\mathbf{U}_1,\mathbf{Y}_3) \leq \frac{1}{n}H(\mathbf{U}_2|W_1,\mathbf{U}_1,\mathbf{Y}_3) \leq \epsilon_{2,n}', \tag{76}$$

$$\frac{1}{n}H(\mathbf{X}|W_1,W_2,\mathbf{U}_1,\mathbf{U}_2,\mathbf{Y}_3) \leq \frac{1}{n}H(\mathbf{X}|W_1,\mathbf{U}_1,\mathbf{U}_2,\mathbf{Y}_3) \leq \epsilon_{1,n}'. \tag{77}$$

Substituting the above into (73), we get

$$H(W_1,W_2|\mathbf{Y}_3) \geq nR_{1e} + nR_{2e} - n\epsilon_{1,2}, \tag{78}$$

where $\epsilon_{1,2} = \delta_1' + 2\delta' + \epsilon_{1,n}' + \epsilon_{2,n}'$, thus satisfying (5c). As a result, we see that the security conditions in (5) are satisfied and we have shown that the rate-equivocation tuple $(R_0, R_1, R_{1e}, R_2, R_{2e})$ is achievable.

V. Outer Bounds for the 3-Receiver BC with Degraded Message Sets

In the derivation of the outer bounds, we note that, for the original Markov chain conditions

$$U_1 \to U_2 \to (U_3, X) \to (Y_1, Y_2, Y_3), \tag{79a}$$

$$U_1 \to U_3 \to (U_2, X) \to (Y_1, Y_2, Y_3), \tag{79b}$$

$$U_1 \to (U_2, U_3) \to X \to (Y_1, Y_2, Y_3), \tag{79c}$$

which arise from the code generation for the 3-receiver BC, there exists the set of conditions

$$U_1 \to \tilde{U}_2 \to U_2 \to (U_3, X) \to (Y_1, Y_2, Y_3), \tag{80a}$$

$$U_1 \to U_3 \to (\tilde{U}_2, U_2, X) \to (Y_1, Y_2, Y_3), \tag{80b}$$

$$U_1 \to (\tilde{U}_2, U_2, U_3) \to X \to (Y_1, Y_2, Y_3), \tag{80c}$$

which come about by inserting auxiliary random variable $\tilde{U}_2$ between U_1 and U_2 in the code generation, so that $U_1 \to \tilde{U}_2 \to U_2$ is satisfied. The code generation and decoding conditions are equivalent for the following:

1) For the 3-receiver BC with 3 degraded message sets, let $\tilde{U}_2$ represent information about W_0, and set $\tilde{U}_2 = U_1$ for equivalent code generation and decoding conditions under (79) and (80);
2) For the 3-receiver BC with 3 degraded message sets, let $\tilde{U}_2$ represent information about W_1, and set $\tilde{U}_2 = U_2$ for equivalent code generation and decoding conditions under both (79) and (80);
3) For the 3-receiver BC with 2 degraded message sets (Type 1), let $\tilde{U}_2$ represent information about W_1, and set $\tilde{U}_2 = U_1$ for equivalent code generation and decoding conditions under (79) and (80).

We will show that case (1) is true in the Appendix of this paper, while cases (2) and (3) are shown to be true in [11, Appendix III].

So, to obtain the outer bound to the rate equivocation region for the 3-receiver BC with 3 degraded message sets, we first find the outer bound $\mathcal{R}'_O$ for the 3-receiver BC using conditions (80). Then we set $\tilde{U}_2 = U_1$ (as in case (1)) to obtain the outer bound to the rate equivocation region for the 3-receiver BC with 3 degraded message sets $\mathcal{R}_O$ with original conditions (79).

For the 3-receiver BC with 2 degraded message sets (Type 1), we use the same procedure.

A. Proof for the 3-receiver BC with 3 degraded message sets

In this section we show the proof for the outer bound in Theorem 2. We use a $(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, n)$-code with error probability $P_e^{(n)}$ with the code construction so that we have the Markov chain condition (W_0, W_1, W_2) $\to \mathbf{X} \to (\mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}_3)$. Then, the probability distribution on $\mathcal{W}_0 \times \mathcal{W}_1 \times \mathcal{W}_2 \times \mathcal{X}^n \times \mathcal{Y}_1^n \times \mathcal{Y}_2^n \times \mathcal{Y}_3^n$ is given by

$$p(w_0)p(w_1)p(w_2)p(\mathbf{x}|w_0, w_1, w_2) \prod_{i=1}^{n} p(y_{1i}, y_{2i}, y_{3i}|x_i). \tag{81}$$

By Fano's inequality, we have

$$\begin{cases} H(W_0|\mathbf{Y}_k) \le nR_0P_e^{(n)}+1 \triangleq n\gamma_k,\ k=1,2,3, \\ H(W_0,W_1|\mathbf{Y}_1) \le n(R_0+R_1)P_e^{(n)}+1 \triangleq n\gamma_4, \\ H(W_0,W_1|\mathbf{Y}_2) \le n(R_0+R_1)P_e^{(n)}+1 \triangleq n\gamma_5, \\ H(W_0,W_2|\mathbf{Y}_1) \le n(R_0+R_2)P_e^{(n)}+1 \triangleq n\gamma_6, \\ H(W_0,W_1,W_2|\mathbf{Y}_1) \le n(R_0+R_1+R_2)P_e^{(n)}+1 \triangleq n\gamma_7, \\ H(W_0,W_2|\mathbf{Y}_2) \le n(R_0+R_2)P_e^{(n)}+1 \triangleq n\gamma_8, \end{cases} \tag{82}$$

where $\gamma_k \to 0$ if $P_e^{(n)} \to 0\ \forall k$. Now we want to define the auxiliary random variables $U_{1,i}$, $\tilde{U}_{2,i}$, $U_{2,i}$, $U_{3,i}$, satisfying the conditions

$$U_{1,i} \to \tilde{U}_{2,i} \to U_{2,i} \to (U_{3,i},X_i) \to (Y_{1,i},Y_{2,i},Y_{3,i}), \tag{83a}$$

$$U_{1,i} \to U_{3,i} \to (\tilde{U}_{2,i},U_{2,i},X_i) \to (Y_{1,i},Y_{2,i},Y_{3,i}), \tag{83b}$$

$$U_{1,i} \to (\tilde{U}_{2,i},U_{2,i},U_{3,i}) \to X_i \to (Y_{1,i},Y_{2,i},Y_{3,i}), \tag{83c}$$

for all i. When we have derived the outer bounds for the rates for conditions (83), we can then set $\tilde{U}_{2,i}=U_{1,i}$ to obtain the rates for the original conditions

$$U_{1,i} \to U_{2,i} \to (U_{3,i},X_i) \to (Y_{1,i},Y_{2,i},Y_{3,i}), \tag{84a}$$

$$U_{1,i} \to U_{3,i} \to (U_{2,i},X_i) \to (Y_{1,i},Y_{2,i},Y_{3,i}), \tag{84b}$$

$$U_{1,i} \to (U_{2,i},U_{3,i}) \to X_i \to (Y_{1,i},Y_{2,i},Y_{3,i}). \tag{84c}$$

Here, however, we will define the auxiliary random variables $U_{1,i} \triangleq (W_0,\mathbf{Y}_1^{i-1})$, $\tilde{U}_{2,i} \triangleq (U_{1,i},\tilde{\mathbf{Y}}_2^{i+1})$, $U_{2,i}=W_1$, $U_{3,i} \triangleq (U_{1,i},\tilde{\mathbf{Y}}_3^{i+1})$ which satisfy the conditions

$$U_{1,i} \to (\tilde{U}_{2,i},U_{2,i}) \to (U_{3,i},X_i) \to (Y_{1,i},Y_{2,i},Y_{3,i}), \tag{85a}$$

$$U_{1,i} \to U_{3,i} \to (\tilde{U}_{2,i},U_{2,i},X_i) \to (Y_{1,i},Y_{2,i},Y_{3,i}), \tag{85b}$$

$$U_{1,i} \to (\tilde{U}_{2,i},U_{2,i},U_{3,i}) \to X_i \to (Y_{1,i},Y_{2,i},Y_{3,i}), \tag{85c}$$

for all i, which are weaker than and included in conditions (83). By setting $\tilde{U}_{2,i}=U_{1,i}$ in (85), we still obtain the original conditions (84). Thus we use (85) in our subsequent derivation for the outer bound.

We first prove three relations which are a consequence of (85).

Relation 1: $I(\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1}) = I(\tilde{\mathbf{Y}}_2^{i+1},\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1}) = I(\tilde{\mathbf{Y}}_2^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1}), k=1,2,3.$

Proof: For any $Y_{k,i}$, $k=1,2,3$, we have

$$\begin{aligned} I(\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1}) &= I(\tilde{\mathbf{Y}}_2^{i+1},\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1}) - I(\tilde{\mathbf{Y}}_2^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) \\ &\overset{(a)}{=} I(\tilde{\mathbf{Y}}_2^{i+1},\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1}) \\ &= I(\tilde{\mathbf{Y}}_2^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1}) + I(\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_2^{i+1}) \\ &\overset{(b)}{=} I(\tilde{\mathbf{Y}}_2^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1}), \end{aligned} \tag{86}$$

where (a) is due to $I(\tilde{\mathbf{Y}}_2^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) = I(\tilde{U}_{2,i};Y_{k,i}|U_{3,i}) = 0$ by (85a) and (b) is due to the fact that $I(\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_2^{i+1}) = I(U_{3,i};Y_{k,i}|\tilde{U}_{2,i}) = 0$ by (85b). ■

Relation 2: $I(\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1}) = I(W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1}) = I(W_1;Y_{k,i}|W_0,\mathbf{Y}_1^{i-1}), k=1,2,3.$

Proof: For any $Y_{k,i}$, $k=1,2,3$, we have

$$\begin{aligned} I(\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1}) &= I(W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1;Y_{k,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) \\ &\overset{(a)}{=} I(W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1}) \\ &= I(W_1;Y_{k,i}|W_0,\mathbf{Y}_1^{i-1}) + I(\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) \\ &\overset{(b)}{=} I(W_1;Y_{k,i}|W_0,\mathbf{Y}_1^{i-1}), \end{aligned} \tag{87}$$

where (a) is due to $I(W_1;Y_{k,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) = I(U_{2,i};Y_{k,i}|U_{3,i}) = 0$ by (85a) and (b) is due to the fact that $I(\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) = I(U_{3,i};Y_{k,i}|U_{2,i},U_{1,i}) = 0$ by (85b). ■

Relation 3: $I(\tilde{\mathbf{Y}}_2^{i+1};Y_{k,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) = I(\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,W_1,\mathbf{Y}_1^{i-1}), k=1,2,3.$

Proof: For any $Y_{k,i}$, $k=1,2,3$, we have

$$\begin{aligned} I(\tilde{\mathbf{Y}}_2^{i+1};Y_{k,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) &= I(\tilde{\mathbf{Y}}_2^{i+1},\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) - I(\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_2^{i+1}) \\ &\overset{(a)}{=} I(\tilde{\mathbf{Y}}_2^{i+1},\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) \\ &= I(\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) + I(\tilde{\mathbf{Y}}_2^{i+1};Y_{k,i}|W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) \\ &\overset{(b)}{=} I(\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,W_1,\mathbf{Y}_1^{i-1}), \end{aligned} \tag{88}$$

where (a) is due to $I(\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_2^{i+1}) = I(U_{3,i};Y_{k,i}|U_{2,i},\tilde{U}_{2,i}) = 0$ by (85b); and (b) is by $I(\tilde{\mathbf{Y}}_2^{i+1};Y_{k,i}|W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) = I(\tilde{U}_{2,i};Y_{k,i}|U_{2,i},U_{3,i}) = H(Y_{k,i}|U_{2,i},U_{3,i}) - H(Y_{k,i}|\tilde{U}_{2,i},U_{2,i},U_{3,i}) = 0$ by (85a). ■

We begin by proving the outer bounds to the equivocation rates. For R_{1e}, we have two possible choices corresponding to whether $\mathbf{X}$ is sent to Y_1 or $\mathbf{U}_2$ is sent to Y_2. For the first case, we have

$$\begin{aligned}
&nR_{1e} \leq H(W_1|\mathbf{Y}_3) + n\tilde{\epsilon}_1 \text{ (by secrecy condition)}\\
&= H(W_1|\mathbf{Y}_3, W_0) + I(W_1;W_0|\mathbf{Y}_3) + n\tilde{\epsilon}_1\\
&= H(W_1|W_0) - I(W_1;\mathbf{Y}_3|W_0) + I(W_1;W_0|\mathbf{Y}_3) + n\tilde{\epsilon}_1\\
&= I(W_1;\mathbf{Y}_1|W_0) + H(W_1|W_0,\mathbf{Y}_1) - I(W_1;\mathbf{Y}_3|W_0) + I(W_1;W_0|\mathbf{Y}_3) + n\tilde{\epsilon}_1\\
&= I(W_1;\mathbf{Y}_1|W_0) - I(W_1;\mathbf{Y}_3|W_0) + H(W_0|\mathbf{Y}_3) - H(W_0|W_1,\mathbf{Y}_3) + H(W_1|W_0,\mathbf{Y}_1) + n\tilde{\epsilon}_1\\
&\leq I(W_1;\mathbf{Y}_1|W_0) - I(W_1;\mathbf{Y}_3|W_0) + H(W_0|\mathbf{Y}_3) + H(W_1|W_0,\mathbf{Y}_1) + n\tilde{\epsilon}_1\\
&\overset{(a)}{\leq} I(W_1;\mathbf{Y}_1|W_0) - I(W_1;\mathbf{Y}_3|W_0) + n(\tilde{\epsilon}_1 + \gamma_3 + \gamma_4), \qquad (89)
\end{aligned}$$

where (a) is by Fano's inequality. Expanding the first two terms of (a) by the chain rule, we obtain

$$I(W_1;\mathbf{Y}_1|W_0) = \sum_{i=1}^{n} I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}), \qquad (90a)$$

$$I(W_1;\mathbf{Y}_3|W_0) = \sum_{i=1}^{n} I(W_1;Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1}). \qquad (90b)$$

Now we have

$$nR_{1e} \leq \sum_{i=1}^{n}\Big[I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1;Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1})\Big] + n(\tilde{\epsilon}_1 + \gamma_3 + \gamma_4). \qquad (91)$$

The terms under the summation can be bounded by

$$\begin{aligned}
&I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1;Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1})\\
&= I(W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1,\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1}) - I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1})\\
&\quad + I(\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,W_1,\tilde{\mathbf{Y}}_3^{i+1})\\
&\overset{(a)}{=} I(W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1,\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1}) - I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1})\\
&\quad + I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1})\\
&= I(W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1,\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1})\\
&= I(X_i,W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(X_i,W_1,\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1})\\
&\quad - [I(X_i;Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) - I(X_i;Y_{3,i}|W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})]\\
&\overset{(b)}{\leq} I(X_i,W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(X_i,W_1,\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1})\\
&\overset{(c)}{=} I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) - I(X_i;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) + I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})\\
&\quad - I(\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1})
\end{aligned}$$

$$\begin{aligned}
&\overset{(d)}{=} I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) - I(X_i;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) + I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) \\
&\quad - I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) \\
&= I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) - I(X_i;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) \\
&= I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) - [I(X_i,\tilde{\mathbf{Y}}_3^{i+1};Y_{3,i}|W_0,\mathbf{Y}_1^{i-1}) - I(\tilde{\mathbf{Y}}_3^{i+1};Y_{3,i}|W_0,\mathbf{Y}_1^{i-1})] \\
&\overset{(e)}{=} I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) - [I(X_i;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1}) - I(\tilde{\mathbf{Y}}_2^{i+1};Y_{3,i}|W_0,\mathbf{Y}_1^{i-1})] \\
&= I(X_i;Y_{1,i}|U_{3,i}) - [I(X_i;Y_{3,i}|U_{1,i}) - I(\tilde{U}_2;Y_{3,i}|U_{1,i})] \overset{(f)}{=} I(X_i;Y_{1,i}|U_{3,i}) - I(X_i;Y_{3,i}|U_{1,i}),
\end{aligned} \tag{92}$$

where (a) has last term by [1, Lemma 7] so that

$$\sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) = \sum_{i=1}^{n} I(\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,W_1,\tilde{\mathbf{Y}}_3^{i+1});$$

(b) is due to $[I(X_i;Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) - I(X_i;Y_{3,i}|W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})] \geq 0$ by the fact that Y_1 is a more capable channel than Y_3 along with the fact that it may be verified using a functional dependency graph [17] that $(W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) \to X_i \to (Y_{1,i},Y_{3,i})$ forms a Markov chain, so the more capable channel condition is satisfied [15]; (c) is because $I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1},X_i) = 0$ by

$$\begin{aligned}
&I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1},X_i) = H(Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1},X_i) - H(Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1},X_i) \\
&= H(Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1},X_i) - H(Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1},X_i) = 0,
\end{aligned} \tag{93}$$

where the second equality is obtained using the relation $W_1 \to X_i \to Y_i$ on the second term on the right-hand-side, and similarly $I(W_1;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1},X_i) = 0$; (d) has last term by [1, Lemma 7] so that

$$\sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) = \sum_{i=1}^{n} I(\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1});$$

(e) has the first term in the square brackets by the fact that $I(\tilde{\mathbf{Y}}_3^{i+1};Y_{3,i}|W_0,\mathbf{Y}_1^{i-1},X_i) = 0$ since given X_i, $\tilde{\mathbf{Y}}_3^{i+1}$ is independent of $Y_{3,i}$ as may be seen using a functional dependency graph, and the second term in the square brackets $I(\tilde{\mathbf{Y}}_3^{i+1};Y_{3,i}|W_0,\mathbf{Y}_1^{i-1}) = I(\tilde{\mathbf{Y}}_2^{i+1};Y_{3,i}|W_0,\mathbf{Y}_1^{i-1})$ by Relation 1; and (f) is by substituting $\tilde{U}_{2,i} = U_{1,i}$.

Then, we have

$$nR_{1e} \leq \sum_{i=1}^{n} [I(X_i;Y_{1,i}|U_{3,i}) - I(X_i;Y_{3,i}|U_{1,i})] + n(\tilde{\epsilon}_1 + \gamma_3 + \gamma_4). \tag{94}$$

Next consider the rate for W_1 sent to receiver Y_2. We have, following (89),

$$\begin{aligned}
nR_{1e} &\leq I(W_1;\mathbf{Y}_2|W_0) - I(W_1;\mathbf{Y}_3|W_0) + H(W_0|\mathbf{Y}_3) + H(W_1|W_0,\mathbf{Y}_2) + n\epsilon_1 \\
&\overset{(a)}{\leq} I(W_1;\mathbf{Y}_2|W_0) - I(W_1;\mathbf{Y}_3|W_0) + n(\epsilon_1 + \gamma_3 + \gamma_5),
\end{aligned} \tag{95}$$

where (a) is by Fano's inequality. For the first two terms in (95), we have

$$\begin{aligned}
&I(W_1;\mathbf{Y}_2|W_0)-I(W_1;\mathbf{Y}_3|W_0)=\sum_{i=1}^{n}\Big[I(W_1;Y_{2,i}|W_0,\tilde{\mathbf{Y}}_2^{i+1})-I(W_1;Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1})\Big]\\
&=\sum_{i=1}^{n}\Big[I(W_1,\mathbf{Y}_1^{i-1};Y_{2,i}|W_0,\tilde{\mathbf{Y}}_2^{i+1})-I(W_1,\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1})-I(\mathbf{Y}_1^{i-1};Y_{2,i}|W_0,W_1,\tilde{\mathbf{Y}}_2^{i+1})\\
&\quad+I(\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,W_1,\tilde{\mathbf{Y}}_3^{i+1})\Big]\\
&\overset{(a)}{=}\sum_{i=1}^{n}\Big[I(W_1,\mathbf{Y}_1^{i-1};Y_{2,i}|W_0,\tilde{\mathbf{Y}}_2^{i+1})-I(W_1,\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1})-I(\tilde{\mathbf{Y}}_2^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1})\\
&\quad+I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1})\Big] \qquad (96)\\
&\overset{(b)}{=}\sum_{i=1}^{n}\Big[I(W_1,\mathbf{Y}_1^{i-1};Y_{2,i}|W_0,\tilde{\mathbf{Y}}_2^{i+1})-I(W_1,\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1})\Big]\\
&=\sum_{i=1}^{n}\Big[I(\mathbf{Y}_1^{i-1};Y_{2,i}|W_0,\tilde{\mathbf{Y}}_2^{i+1})+I(W_1;Y_{2,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_2^{i+1})-I(X_i,W_1,\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1})\\
&\quad+I(X_i;Y_{3,i}|W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})\Big]\\
&\overset{(c)}{\le}\sum_{i=1}^{n}\Big[I(\tilde{\mathbf{Y}}_2^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})+I(W_1;Y_{2,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_2^{i+1})-I(\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1})\\
&\quad-I(X_i;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})+I(X_i;Y_{3,i}|W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})\Big] \qquad (97)\\
&\overset{(d)}{=}\sum_{i=1}^{n}\Big[I(\tilde{\mathbf{Y}}_2^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})+I(W_1;Y_{2,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_2^{i+1})-I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})\\
&\quad-I(X_i;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1})+I(\tilde{\mathbf{Y}}_3^{i+1};Y_{3,i}|W_0,\mathbf{Y}_1^{i-1})+I(X_i;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1})-I(W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{3,i}|W_0,\mathbf{Y}_1^{i-1})\Big]\\
&\overset{(e)}{\le}\sum_{i=1}^{n}\Big[I(W_1;Y_{2,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_2^{i+1})+I(\tilde{\mathbf{Y}}_2^{i+1};Y_{3,i}|W_0,\mathbf{Y}_1^{i-1})-I(W_1;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1})\Big]\\
&=\sum_{i=1}^{n}\Big[I(U_{2,i};Y_{2,i}|\tilde{U}_{2,i},U_{1,i})+I(\tilde{U}_{2,i};Y_{3,i}|U_{1,i})-I(U_{2,i};Y_{3,i}|U_{1,i})\Big]=\sum_{i=1}^{n}[I(U_{2,i};Y_{2,i}|U_{1,i})-I(U_{2,i};Y_{3,i}|U_{1,i})]\\
&\qquad (98)
\end{aligned}$$

where (a) has the last two terms by [1, Lemma 7] which gives

$$\begin{aligned}
\sum_{i=1}^{n}I(\mathbf{Y}_1^{i-1};Y_{2,i}|W_0,W_1,\tilde{\mathbf{Y}}_2^{i+1})&=\sum_{i=1}^{n}I(\tilde{\mathbf{Y}}_2^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1}),\\
\sum_{i=1}^{n}I(\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,W_1,\tilde{\mathbf{Y}}_3^{i+1})&=\sum_{i=1}^{n}I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1});
\end{aligned}$$

(b) is because the last two terms in (96) above are $I(\tilde{\mathbf{Y}}_2^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1})=I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1})=I(U_{3,i};Y_{1,i}|U_{2,i},U_{1,i})=0$ by Relation 3 and (85b) and similarly $I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1})=0$; (c) is by [1,

Lemma 7] which gives

$$\sum_{i=1}^{n} I(\mathbf{Y}_1^{i-1}; Y_{2,i}|W_0, \tilde{\mathbf{Y}}_2^{i+1}) = \sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_2^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1});$$

(d) is by [1, Lemma 7] from which

$$\sum_{i=1}^{n} I(\mathbf{Y}_1^{i-1}; Y_{3,i}|W_0, \tilde{\mathbf{Y}}_3^{i+1}) = \sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_3^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1});$$

and (e) is by $I(\tilde{\mathbf{Y}}_2^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) = I(\tilde{\mathbf{Y}}_3^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1})$ and $I(\tilde{\mathbf{Y}}_3^{i+1}; Y_{3,i}|W_0, \mathbf{Y}_1^{i-1}) = I(\tilde{\mathbf{Y}}_2^{i+1}; Y_{3,i}|W_0, \mathbf{Y}_1^{i-1})$ by Relation 1. Consequently we have

$$nR_{1e} \le \sum_{i=1}^{n} [I(U_{2,i}; Y_{2,i}|U_{1,i}) - I(U_{2,i}; Y_{3,i}|U_{1,i})] + n(\epsilon_1 + \gamma_3 + \gamma_5). \tag{99}$$

For the equivocation rate R_{2e}, we consider W_2 sent to receiver Y_1 using codeword $\mathbf{X}$. Following the same procedure to obtain (89), we get

$$\begin{aligned} nR_{2e} &\le H(W_2|\mathbf{Y}_3) + n\epsilon_2 \\ &\le I(W_2; \mathbf{Y}_1|W_0) - I(W_2; \mathbf{Y}_3|W_0) + H(W_0|\mathbf{Y}_3) + H(W_2|W_0, \mathbf{Y}_1) \\ &\le I(W_2; \mathbf{Y}_1|W_0) - I(W_2; \mathbf{Y}_3|W_0) + n(\epsilon_2 + \gamma_3 + \gamma_6), \end{aligned} \tag{100}$$

by Fano's inequality. Expanding the first two terms of the inequality above by the chain rule and following the same procedure as for R_{1e} in (90a), (90b) to (91), we obtain

$$nR_{2e} \le \sum_{i=1}^{n} \left[I(W_2; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) - I(W_2; Y_{3,i}|W_0, \tilde{\mathbf{Y}}_3^{i+1}) \right] + n(\epsilon_2 + \gamma_3 + \gamma_6). \tag{101}$$

The terms under the summation can be bounded as

$$\begin{aligned} & I(W_2; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) - I(W_2; Y_{3,i}|W_0, \tilde{\mathbf{Y}}_3^{i+1}) \\ & = I(\tilde{\mathbf{Y}}_3^{i+1}, W_2; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) - I(W_2, \mathbf{Y}_1^{i-1}; Y_{3,i}|W_0, \tilde{\mathbf{Y}}_3^{i+1}) - I(\tilde{\mathbf{Y}}_3^{i+1}; Y_{1,i}|W_0, W_2, \mathbf{Y}_1^{i-1}) \\ & \quad + I(\mathbf{Y}_1^{i-1}; Y_{3,i}|W_0, W_2, \tilde{\mathbf{Y}}_3^{i+1}) \\ & \overset{(a)}{=} I(\tilde{\mathbf{Y}}_3^{i+1}, W_2; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) - I(W_2, \mathbf{Y}_1^{i-1}; Y_{3,i}|W_0, \tilde{\mathbf{Y}}_3^{i+1}) - I(\tilde{\mathbf{Y}}_3^{i+1}; Y_{1,i}|W_0, W_2, \mathbf{Y}_1^{i-1}) \\ & \quad + I(\tilde{\mathbf{Y}}_3^{i+1}; Y_{1,i}|W_0, W_2, \mathbf{Y}_1^{i-1}) \\ & = I(\tilde{\mathbf{Y}}_3^{i+1}, W_2; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) - I(W_2, \mathbf{Y}_1^{i-1}; Y_{3,i}|W_0, \tilde{\mathbf{Y}}_3^{i+1}) \\ & = I(\tilde{\mathbf{Y}}_3^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) + I(W_2; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_3^{i+1}) - I(\mathbf{Y}_1^{i-1}; Y_{3,i}|W_0, \tilde{\mathbf{Y}}_3^{i+1}) \\ & \quad - I(W_2; Y_{3,i}|W_0, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_3^{i+1}) \\ & \overset{(b)}{=} I(\tilde{\mathbf{Y}}_3^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) + I(W_2; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_3^{i+1}) - I(\tilde{\mathbf{Y}}_3^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) \\ & \quad - I(W_2; Y_{3,i}|W_0, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_3^{i+1}) \end{aligned}$$

$$
\begin{aligned}
&= I(W_2; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_3^{i+1}) - I(W_2; Y_{3,i}|W_0, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_3^{i+1}) && \text{(102a)}\\
&= I(X_i, W_2; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_3^{i+1}) - I(X_i, W_2; Y_{3,i}|W_0, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_3^{i+1})\\
&\quad - [I(X_i; Y_{1,i}|W_0, W_2, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_3^{i+1}) - I(X_i; Y_{3,i}|W_0, W_2, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_3^{i+1})]\\
&\overset{(c)}{\leq} I(X_i, W_2; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_3^{i+1}) - I(X_i, W_2; Y_{3,i}|W_0, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_3^{i+1})\\
&= I(X_i, W_2, \tilde{\mathbf{Y}}_3^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) - I(X_i, W_2, \tilde{\mathbf{Y}}_3^{i+1}; Y_{3,i}|W_0, \mathbf{Y}_1^{i-1})\\
&\quad - [I(\tilde{\mathbf{Y}}_3^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) - I(\tilde{\mathbf{Y}}_3^{i+1}; Y_{3,i}|W_0, \mathbf{Y}_1^{i-1})]\\
&\overset{(d)}{\leq} I(X_i; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) - I(X_i; Y_{3,i}|W_0, \mathbf{Y}_1^{i-1}) - [I(W_1; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) - I(W_1; Y_{3,i}|W_0, \mathbf{Y}_1^{i-1})]\\
&= I(X_i; Y_{1,i}|U_{2,i}) - I(X_i; Y_{3,i}|U_{2,i}) && \text{(102b)}
\end{aligned}
$$

where (a) is by [1, Lemma 7] so that

$$\sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_3^{i+1}; Y_{1,i}|W_0, W_2, \mathbf{Y}_1^{i-1}) = \sum_{i=1}^{n} I(\mathbf{Y}_1^{i-1}; Y_{3,i}|W_0, W_2, \tilde{\mathbf{Y}}_3^{i+1}),$$

(b) is also by [1, Lemma 7] giving

$$\sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_3^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) = \sum_{i=1}^{n} I(\mathbf{Y}_1^{i-1}; Y_{3,i}|W_0, \tilde{\mathbf{Y}}_3^{i+1}),$$

(c) is because $[I(X_i; Y_{1,i}|W_0, W_2, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_3^{i+1}) - I(X_i; Y_{3,i}|W_0, W_2, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_3^{i+1})] \geq 0$ since Y_1 is a more capable channel than Y_3 and $(W_0, W_2, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_3^{i+1}) \to X_i \to (Y_{1,i}, Y_{3,i})$ forms a Markov chain so satisfying the more capable channel condition, and (d) is due to firstly,

$$
\begin{aligned}
&I(W_2, \tilde{\mathbf{Y}}_3^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}, X_i) = H(Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}, X_i) - H(Y_{1,i}|W_0, W_2, \tilde{\mathbf{Y}}_3^{i+1}, \mathbf{Y}_1^{i-1}, X_i)\\
&\quad = H(Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}, X_i) - H(Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}, X_i) = 0, && \text{(103)}
\end{aligned}
$$

which is true since, given X_i, $\tilde{\mathbf{Y}}_3^{i+1}$ is independent of $Y_{3,i}$ as can be verified using a functional dependency graph, and by $W_2 \to X_i \to Y_{1,i}$; secondly, $I(W_2, \tilde{\mathbf{Y}}_3^{i+1}; Y_{3,i}|W_0, \mathbf{Y}_1^{i-1}, X_i) = 0$ since given X_i, $\tilde{\mathbf{Y}}_3^{i+1}$ is independent of $Y_{3,i}$ and $W_2 \to X_i \to Y_{3,i}$; and thirdly, $I(\tilde{\mathbf{Y}}_3^{i+1}; Y_{k,i}|W_0, \mathbf{Y}_1^{i-1}) = I(W_1; Y_{k,i}|W_0, \mathbf{Y}_1^{i-1})$ for $k = 1, 3$ from Relation 2. Then, we shall have

$$nR_{2e} \leq \sum_{i=1}^{n} [I(X_i; Y_{1,i}|U_{2,i}) - I(X_i; Y_{3,i}|U_{2,i})] + n(\epsilon_2 + \gamma_3 + \gamma_6). \tag{104}$$

For the rates $(R_{1e} + R_{2e})$, consider the combined message (W_1, W_2) sent to receiver Y_1 using codeword $\mathbf{X}$.

It can be shown that

$$\begin{aligned}
&n(R_{1e}+R_{2e}) \leq H(W_1,W_2|\mathbf{Y}_3)+\epsilon_{1,2}\\
&\overset{(a)}{\leq} I(W_1,W_2;\mathbf{Y}_1|W_0)-I(W_1,W_2;\mathbf{Y}_3|W_0)+H(W_0|\mathbf{Y}_3)+H(W_1,W_2|W_0,\mathbf{Y}_1)+\epsilon_{1,2}\\
&\overset{(b)}{\leq} I(W_1,W_2;\mathbf{Y}_1|W_0)-I(W_1,W_2;\mathbf{Y}_3|W_0)+n(\epsilon_{1,2}+\gamma_3+\gamma_7)\\
&\overset{(c)}{\leq} \sum_{i=1}^{n}\Big[I(W_1,W_2;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})-I(W_1,W_2;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})\Big]+n(\epsilon_{1,2}+\gamma_3+\gamma_7) \qquad (105)
\end{aligned}$$

where (a) results in following the steps in (89) using (W_1,W_2) instead of W_1, (b) is by Fano's inequality, and (c) results in following the steps to obtain (102a) using (W_1,W_2) instead of W_2. The terms under the summation can be bounded as

$$\begin{aligned}
&I(W_1,W_2;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})-I(W_1,W_2;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})\\
&=I(X_i,W_1,W_2;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})-I(X_i,W_1,W_2;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})\\
&\quad-[I(X_i;Y_{1,i}|W_0,W_1,W_2,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})-I(X_i;Y_{3,i}|W_0,W_1,W_2,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})]\\
&\overset{(a)}{\leq} I(X_i,W_1,W_2;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})-I(X_i,W_1,W_2;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})\\
&=I(X_i,W_1,W_2,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})-I(X_i,W_1,W_2,\tilde{\mathbf{Y}}_3^{i+1};Y_{3,i}|W_0,\mathbf{Y}_1^{i-1})\\
&\quad-[I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})-I(\tilde{\mathbf{Y}}_3^{i+1};Y_{3,i}|W_0,\mathbf{Y}_1^{i-1})]\\
&\overset{(b)}{=} I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})-I(X_i;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1})-[I(\tilde{\mathbf{Y}}_2^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})-I(\tilde{\mathbf{Y}}_2^{i+1};Y_{3,i}|W_0,\mathbf{Y}_1^{i-1})]\\
&=I(X_i;Y_{1,i}|U_{1,i})-I(X_i;Y_{3,i}|U_{1,i})-[I(\tilde{U}_{2,i};Y_{1,i}|U_{1,i})-I(\tilde{U}_{2,i};Y_{3,i}|U_{1,i})]\\
&\overset{(c)}{=} I(X_i;Y_{1,i}|U_{1,i})-I(X_i;Y_{3,i}|U_{1,i}), \qquad (106)
\end{aligned}$$

where (a) is due to Y_1 being a more capable channel than Y_1 which gives $[I(X_i;Y_{1,i}|W_0,W_1,W_2,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})-I(X_i;Y_{3,i}|W_0,W_1,W_2,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})]\geq 0$ for $(W_0,W_1,W_2,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1})\rightarrow X_i\rightarrow(Y_{1,i},Y_{3,i})$ as may be verified using a functional dependency graph; (b) is due to, first, that

$$\begin{aligned}
&I(W_1,W_2,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},X_i)=H(Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},X_i)-H(Y_{1,i}|W_0,W_1,W_2,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1},X_i)\\
&\;=H(Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},X_i)-H(Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},X_i)=0 \qquad (107)
\end{aligned}$$

since in the second term in the second equality is obtained using the relation $(W_1,W_2)\rightarrow X_i\rightarrow Y_{1,i}$ and the fact that given X_i, $\tilde{\mathbf{Y}}_3^{i+1}$ is independent of $Y_{1,i}$, secondly, we can obtain $I(W_1,W_2;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1},X_i)=0$ in a similar way, and thirdly we have, by Relation 1, $I(\tilde{\mathbf{Y}}_3^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1})=I(\tilde{\mathbf{Y}}_2^{i+1};Y_{k,i}|W_0,\mathbf{Y}_1^{i-1})$, $k=1,3$;

(c) is by substituting $\tilde{U}_{2,i} = U_{1,i}$. Then, we have

$$n(R_{1e}+R_{2e}) \le \sum_{i=1}^{n} [I(X_i;Y_{1,i}|U_{1,i}) - I(X_i;Y_{3,i}|U_{1,i}),] + n(\epsilon_{1,2}+\gamma_3+\gamma_7). \qquad (108)$$

We now prove the rates for R_0, R_0+R_1, R_0+R_2 and $R_0+R_1+R_2$. For rate R_0, we have

$$\begin{aligned}
nR_0 &= H(W_0) = I(W_0;\mathbf{Y}_1) + H(W_0|\mathbf{Y}_1) \\
&\le I(W_0;\mathbf{Y}_1) + n\gamma_1 \qquad \text{by Fano's inequality} \\
&= \sum_{i=1}^{n} I(W_0;Y_{1,i}|\mathbf{Y}_1^{i-1}) + n\gamma_1 \\
&\le \sum_{i=1}^{n} I(W_0,\mathbf{Y}_1^{i-1};Y_{1,i}) + n\gamma_1 && (109) \\
&= \sum_{i=1}^{n} I(U_{1,i};Y_{1,i}) + n\gamma_1. && (110)
\end{aligned}$$

We also have

$$\begin{aligned}
nR_0 &= H(W_0) = I(W_0;\mathbf{Y}_3) + H(W_0|\mathbf{Y}_3) \\
&\le I(W_0;\mathbf{Y}_3) + n\gamma_3 \qquad \text{by Fano}'\text{s inequality} \\
&= \sum_{i=1}^{n} I(W_0;Y_{3,i}|\tilde{\mathbf{Y}}_3^{i+1}) + n\gamma_3 \\
&= \sum_{i=1}^{n} \Big[I(W_0,\mathbf{Y}_1^{i-1};Y_{3,i}|\tilde{\mathbf{Y}}_3^{i+1}) - I(\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1})\Big] + n\gamma_3 \\
&\overset{(a)}{\le} \sum_{i=1}^{n} \Big[I(W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1};Y_{3,i}) - I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})\Big] + n\gamma_3 && (111) \\
&= \sum_{i=1}^{n} [I(U_{3,i};Y_{3,i}) - I(U_{3,i};Y_{1,i}|U_{1,i})] + n\gamma_3 && (112)
\end{aligned}$$

where (a) is by [1,Lemma 7] from which $\sum_{i=1}^{n} I(\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1}) = \sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})$.

For the rates (R_0+R_1), we consider the following cases when the messages are sent:

1) Case 1: W_1 sent to Y_1, W_0 sent to Y_1 or Y_3;
2) Case 2: W_1 sent to Y_2, W_0 sent to Y_1 or Y_3;
3) Case 3: W_0, W_1 both sent to Y_2.

For Case 1, we have

$$\begin{aligned}
n(R_0+R_1) &= H(W_0,W_1) = H(W_0) + H(W_1|W_0) \\
&= H(W_0) + I(W_1;\mathbf{Y}_1|W_0) + H(W_1|W_0,\mathbf{Y}_1) \\
&\overset{(a)}{\le} H(W_0) + \sum_{i=1}^{n} I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) + \gamma_4
\end{aligned}$$

where (a) is by expanding using the chain rule and using Fano's inequality. Then, on combining with $H(W_0)$ using (109), we can get

$$\begin{aligned} n(R_0+R_1) &\leq \sum_{i=1}^{n} \Big[I(W_0, \mathbf{Y}_1^{i-1}; Y_{1,i}) + I(W_1; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) \Big] + n(\gamma_1+\gamma_4) \\ &= \sum_{i=1}^{n} I(W_1; Y_{1,i}) + n(\gamma_1+\gamma_4) = \sum_{i=1}^{n} I(U_{2,i}; Y_{1,i}) + n(\gamma_1+\gamma_4) \end{aligned} \tag{113}$$

and, combining with $H(W_0)$ using (111), we obtain

$$\begin{aligned} n(R_0+R_1) &\leq \sum_{i=1}^{n} \Big[I(W_0, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_3^{i+1}; Y_{3,i}) - I(\tilde{\mathbf{Y}}_3^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) + I(W_1; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) \Big] + n(\gamma_3+\gamma_4) \\ &\leq \sum_{i=1}^{n} [I(U_{3,i}; Y_{3,i}) + I(U_{2,i}; Y_{1,i}|U_{1,i})] + n(\gamma_3+\gamma_4). \end{aligned} \tag{114}$$

For Case 2, we similarly have

$$\begin{aligned} & n(R_0+R_1) = H(W_0) + I(W_1; \mathbf{Y}_2|W_0) + H(W_1|W_0, \mathbf{Y}_2) \\ & \leq H(W_0) + \sum_{i=1}^{n} I(W_1; Y_{2,i}|W_0, \tilde{\mathbf{Y}}_2^{i+1}) + n\gamma_5 \\ & = H(W_0) + \sum_{i=1}^{n} \Big[I(W_1, \mathbf{Y}_1^{i-1}; Y_{2,i}|W_0, \tilde{\mathbf{Y}}_2^{i+1}) - I(\mathbf{Y}_1^{i-1}; Y_{2,i}|W_0, W_1, \tilde{\mathbf{Y}}_2^{i+1}) \Big] + n\gamma_5 \\ & \overset{(a)}{=} H(W_0) + \sum_{i=1}^{n} \Big[I(\mathbf{Y}_1^{i-1}; Y_{2,i}|W_0, \tilde{\mathbf{Y}}_2^{i+1}) + I(W_1; Y_{2,i}|W_0, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_2^{i+1}) - I(\tilde{\mathbf{Y}}_2^{i+1}; Y_{1,i}|W_0, W_1, \mathbf{Y}_1^{i-1}) \Big] + n\gamma_5 \\ & \overset{(b)}{=} H(W_0) + \sum_{i=1}^{n} \Big[I(\tilde{\mathbf{Y}}_2^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}) + I(W_1; Y_{2,i}|W_0, \mathbf{Y}_1^{i-1}, \tilde{\mathbf{Y}}_2^{i+1}) \Big] + n\gamma_5 \\ & = H(W_0) + \sum_{i=1}^{n} \Big[I(\tilde{U}_{2,i}; Y_{1,i}|U_{1,i}) + I(U_{2,i}; Y_{2,i}|\tilde{U}_{2,i}, U_{1,i}) \Big] + n\gamma_5 = H(W_0) + \sum_{i=1}^{n} I(U_{2,i}; Y_{2,i}|U_{1,i}) + n\gamma_5 \end{aligned} \tag{115}$$

where (a) has the last term in the sum by [1, Lemma 7] giving

$$\sum_{i=1}^{n} I(\mathbf{Y}_1^{i-1}; Y_{2,i}|W_0, W_1, \tilde{\mathbf{Y}}_2^{i+1}) = \sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_2^{i+1}; Y_{1,i}|W_0, W_1, \mathbf{Y}_1^{i-1});$$

(b) is by $I(\tilde{\mathbf{Y}}_2^{i+1}; Y_{1,i}|W_0, W_1, \mathbf{Y}_1^{i-1}) = I(\tilde{\mathbf{Y}}_3^{i+1}; Y_{1,i}|W_0, W_1, \mathbf{Y}_1^{i-1}) = I(U_{3,i}; Y_{1,i}|U_{2,i}, U_{1,i}) = 0$ from Relation 3 and (85b) and first term under the sum by [1, Lemma 7] from which

$$\sum_{i=1}^{n} I(\mathbf{Y}_1^{i-1}; Y_{2,i}|W_0, \tilde{\mathbf{Y}}_2^{i+1}) = \sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_2^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_1^{i-1}).$$

Combining with $H(W_0)$ using (109) and (111), we obtain

$$n(R_0+R_1) \leq \sum_{i=1}^{n} [I(U_{1,i}; Y_{1,i}) + I(U_{2,i}; Y_{2,i}|U_{1,i})] + n(\gamma_1+\gamma_5), \tag{116}$$

$$n(R_0+R_1) \leq \sum_{i=1}^{n} [I(U_{3,i}; Y_{3,i}) + I(U_{2,i}; Y_{2,i}|U_{1,i})] + n(\gamma_3+\gamma_5). \tag{117}$$

For Case 3 we have

$$\begin{aligned}
n(R_0+R_1) &= H(W_0,W_1) = I(W_0,W_1;\mathbf{Y}_2) + H(W_0,W_1|\mathbf{Y}_2) \\
&\overset{(a)}{\leq} \sum_{i=1}^{n} I(W_0,W_1;Y_{2,i}|\tilde{\mathbf{Y}}_2^{i+1}) + n\gamma_5 \\
&= \sum_{i=1}^{n} \Big[I(W_0,W_1,\mathbf{Y}_1^{i-1};Y_{2,i}|\tilde{\mathbf{Y}}_2^{i+1}) - I(\mathbf{Y}_1^{i-1};Y_{2,i}|W_0,W_1,\tilde{\mathbf{Y}}_2^{i+1}) \Big] + n\gamma_5 \\
&\overset{(b)}{=} \sum_{i=1}^{n} \Big[I(W_0,W_1,\mathbf{Y}_1^{i-1};Y_{2,i}|\tilde{\mathbf{Y}}_2^{i+1}) - I(\tilde{\mathbf{Y}}_2^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) \Big] + n\gamma_5 \\
&\leq \sum_{i=1}^{n} I(\tilde{U}_{2,i},U_{2,i};Y_{2,i}) + n\gamma_5 = \sum_{i=1}^{n} I(U_{2,i};Y_{2,i}) + n\gamma_5. \qquad (118)
\end{aligned}$$

where (a) is by Fano's inequality, and (b) has second term in the sum by [1, Lemma 7].

For rates (R_0+R_2) consider message W_2 sent to receiver Y_1 and W_0 sent to either Y_1 or Y_3. To begin, we have

$$\begin{aligned}
& n(R_0+R_2) = H(W_0) + H(W_2|W_0) = H(W_0) + I(W_2;\mathbf{Y}_1|W_0) + H(W_2;\mathbf{Y}_1|W_0) \\
& \overset{(a)}{\leq} H(W_0) + \sum_{i=1}^{n} I(W_2;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) + n\gamma_6 \leq H(W_0) + \sum_{i=1}^{n} I(W_2;W_1,Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) + n\gamma_6 \\
& \overset{(b)}{=} H(W_0) + \sum_{i=1}^{n} I(W_2;Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) + n\gamma_6 \overset{(c)}{\leq} H(W_0) + \sum_{i=1}^{n} I(X_i;Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) + n\gamma_6 \\
& = H(W_0) + \sum_{i=1}^{n} \Big[I(X_i,W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) \Big] + n\gamma_6 \qquad (119) \\
& \overset{(d)}{=} H(W_0) + \sum_{i=1}^{n} \Big[I(X_i,W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) \Big] + n\gamma_6 \\
& \overset{(e)}{=} H(W_0) + \sum_{i=1}^{n} \Big[I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) \Big] + n\gamma_6 \\
& = H(W_0) + \sum_{i=1}^{n} I(X_i;Y_{1,i}|U_{2,i},U_{3,i}) + n\gamma_6 \qquad (120)
\end{aligned}$$

where (a) is by Fano's inequality; (b) is by the independence of W_1 and W_2; (c) is by $W_1 \to X_i \to Y_{1,i}$; (d) is by Relation 2; and (e) is because

$$\begin{aligned}
I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},X_i) &= H(Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},X_i) - H(Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1},X_i) \\
&= H(Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},X_i) - H(Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},X_i) = 0 \qquad (121)
\end{aligned}$$

with the second term in the second equality being due to $W_1 \to X_i \to Y_{1,i}$. Combine the results with $H(W_0)$ in two ways. Firstly, we do this by combining with (110) using (120) to get

$$n(R_0+R_2) \leq \sum_{i=1}^{n} [I(U_{1,i};Y_{1,i}) + I(X_i;Y_{1,i}|U_{2,i},U_{3,i})] + n(\gamma_1+\gamma_6). \qquad (122)$$

Next combine with (111) using (119) to get

$$\begin{aligned}
n(R_0+R_2) &\leq \sum_{i=1}^{n} \Big[I(W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1};Y_{3,i}) - I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) + I(X_i,W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) \\
&\quad - I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) \Big] + n(\gamma_3+\gamma_6) \\
&= \sum_{i=1}^{n} \left[I(U_{3,i};Y_{3,i}) - I(U_{3,i};Y_{1,i}|U_{1,i}) + I(X_i;Y_{1,i}|U_{1,i}) - I(U_{2,i};Y_{1,i}|U_{1,i}) \right] + n(\gamma_3+\gamma_6) \\
&\overset{(a)}{\leq} \sum_{i=1}^{n} \left[I(U_{3,i};Y_{3,i}) + I(X_i;Y_{1,i}|U_{1,i}) - I(U_{2,i},U_{3,i};Y_{1,i}|U_{1,i}) \right] + n(\gamma_3+\gamma_6) \\
&= \sum_{i=1}^{n} \left[I(U_{3,i};Y_{3,i}) + I(X_i;Y_{1,i}|U_{2,i},U_{3,i}) \right] + n(\gamma_3+\gamma_6), \qquad (123)
\end{aligned}$$

where (a) is by

$$\begin{aligned}
&I(U_{2,i};Y_{1,i}|U_{1,i}) + I(U_{3,i};Y_{1,i}|U_{1,i}) = I(U_{2,i},U_{3,i};Y_{1,i}|U_{1,i}) - I(U_{3,i};Y_{1,i}|U_{1,i},U_{2,i}) + I(U_{3,i};Y_{1,i}|U_{1,i}) \\
&\geq I(U_{2,i},U_{3,i};Y_{1,i}|U_{1,i}) - I(U_{3,i};Y_{1,i}|U_{1,i},U_{2,i}) + I(U_{3,i};Y_{1,i}|U_{1,i},U_{2,i}) = I(U_{2,i},U_{3,i};Y_{1,i}|U_{1,i}) \qquad (124)
\end{aligned}$$

with the inequality obtained using (85a).

Lastly, for the rates $(R_0+R_1+R_2)$, consider the following combinations of messages sent to the receivers:

1) Case 1: W_1,W_2 sent to Y_1, W_0 sent to Y_1 or Y_3,
2) Case 2: W_1 sent to Y_2, W_2 sent to Y_1, W_0 sent to Y_1 or Y_3,
3) Case 3: W_0,W_1 sent to Y_2, W_2 sent to Y_1.

For Case 1, we have

$$\begin{aligned}
&n(R_0+R_1+R_2) = H(W_0) + H(W_1,W_2|W_0) = H(W_0) + I(W_1,W_2;\mathbf{Y}_1|W_0) + H(W_1,W_2|W_0,\mathbf{Y}_1) \\
&\leq H(W_0) + \sum_{i=1}^{n} \Big[I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) + I(W_2;Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) \Big] + n\gamma_7 \\
&\overset{(a)}{\leq} H(W_0) + \sum_{i=1}^{n} \Big[I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) + I(X_i;Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) \Big] + n\gamma_7 \\
&= H(W_0) + \sum_{i=1}^{n} I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) + n\gamma_7, \qquad (125)
\end{aligned}$$

where (a) is by $W_1 \to X_i \to Y_{1,i}$. Then combining with (109), we have

$$n(R_0+R_1+R_2) \leq \sum_{i=1}^{n} \Big[I(W_0,\mathbf{Y}_1^{i-1};Y_{1,i}) + I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) \Big] + n(\gamma_1+\gamma_7) = \sum_{i=1}^{n} I(X_i;Y_{1,i}) + n(\gamma_1+\gamma_7). \qquad (126)$$

Combining (125) with (111), we have

$$\begin{aligned} n(R_0+R_1+R_2) &\le \sum_{i=1}^{n} \Big[I(W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1};Y_{3,i}) - I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) + I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})\Big] + n(\gamma_3+\gamma_7) \\ &= \sum_{i=1}^{n} [I(U_{3,i};Y_{3,i}) + I(X_i;Y_{1,i}|U_{3,i})] + n(\gamma_3+\gamma_7). \end{aligned} \tag{127}$$

For Case 2 we have

$$\begin{aligned} & n(R_0+R_1+R_2) = H(W_0) + H(W_1|W_0) + H(W_2|W_0,W_1) \\ & = H(W_0) + I(W_1;\mathbf{Y}_2|W_0) + H(W_2|W_0,\mathbf{Y}_2) + I(W_2;\mathbf{Y}_1|W_0,W_1) + H(W_2|W_0,W_1,\mathbf{Y}_1) \\ & \le H(W_0) + \sum_{i=1}^{n} \Big[I(W_1;Y_{2,i}|W_0,\tilde{\mathbf{Y}}_2^{i+1}) + I(W_2;Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1})\Big] + n(\gamma_5+\gamma_8) \\ & \overset{(a)}{\le} H(W_0) + \sum_{i=1}^{n} \Big[I(W_1,\mathbf{Y}_1^{i-1};Y_{2,i}|W_0,\tilde{\mathbf{Y}}_2^{i+1}) - I(\mathbf{Y}_1^{i-1};Y_{2,i}|W_0,W_1,\tilde{\mathbf{Y}}_2^{i+1}) \\ & \quad +I(X_i;Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1})\Big] + n(\gamma_5+\gamma_8) \\ & = H(W_0) + \sum_{i=1}^{n} \Big[I(\mathbf{Y}_1^{i-1};Y_{2,i}|W_0,\tilde{\mathbf{Y}}_2^{i+1}) + I(W_1;Y_{2,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_2^{i+1}) - I(\mathbf{Y}_1^{i-1};Y_{2,i}|W_0,W_1,\tilde{\mathbf{Y}}_2^{i+1}) \\ & \quad +I(X_i,W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})\Big] + n(\gamma_5+\gamma_8) \\ & \overset{(b)}{=} H(W_0) + \sum_{i=1}^{n} \Big[I(\tilde{\mathbf{Y}}_2^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) + I(W_1;Y_{2,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_2^{i+1}) - I(\tilde{\mathbf{Y}}_2^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) \\ & \quad +I(X_i,W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) + I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1})\Big] + n(\gamma_5+\gamma_8) \\ & \overset{(c)}{=} H(W_0) + \sum_{i=1}^{n} \Big[I(\tilde{\mathbf{Y}}_2^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) + I(W_1;Y_{2,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_2^{i+1}) \\ & \quad +I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})\Big] + n(\gamma_5+\gamma_8) \end{aligned} \tag{128}$$

$$\begin{aligned} & = H(W_0) + \sum_{i=1}^{n} \Big[I(\tilde{U}_{2,i};Y_{1,i}|U_{1,i}) + I(U_{2,i};Y_{2,i}|\tilde{U}_{2,i},U_{1,i}) + I(X_i;Y_{1,i}|U_{2,i},U_{3,i})\Big] + n(\gamma_5+\gamma_8) \\ & = H(W_0) + \sum_{i=1}^{n} [I(U_{2,i};Y_{2,i}|U_{1,i}) + I(X_i;Y_{1,i}|U_{2,i},U_{3,i})] + n(\gamma_5+\gamma_8) \end{aligned} \tag{129}$$

where (a) is by $W_2 \to X_i \to Y_{1,i}$; (b) is by [1, Lemma 7] which gives

$$\sum_{i=1}^{n} I(\mathbf{Y}_1^{i-1};Y_{2,i}|W_0,\tilde{\mathbf{Y}}_2^{i+1}) = \sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_2^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}),$$
$$\sum_{i=1}^{n} I(\mathbf{Y}_1^{i-1};Y_{2,i}|W_0,W_1,\tilde{\mathbf{Y}}_2^{i+1}) = \sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_2^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1});$$

(c) is by $I(\tilde{\mathbf{Y}}_2^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) = I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) = I(U_{3,i};Y_{1,i}|U_{2,i},U_{1,i}) = 0$ from Relation 3 and (85b), and also we have $I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},X_i) = 0$ from $W_1 \to X_i \to Y_{1,i}$ and $I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1})= 0$ from (85b).

Now combine (109) with (129) to get

$$n(R_0+R_1+R_2) \leq \sum_{i=1}^{n} [I(U_{1,i};Y_{1,i}) + I(U_{2,i};Y_{2,i}|U_{1,i}) + I(X_i;Y_{1,i}|U_{2,i},U_{3,i})] + n(\gamma_1+\gamma_5+\gamma_8). \quad (130)$$

Next combine (111) with (128), so that

$$\begin{aligned}
&n(R_0+R_1+R_2)\\
&\leq \sum_{i=1}^{n} \Big[I(W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1};Y_{3,i}) - I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) + I(\tilde{\mathbf{Y}}_2^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})\\
&\quad +I(W_1;Y_{2,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_2^{i+1}) + I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})\Big] + n(\gamma_3+\gamma_5+\gamma_8)\\
&\overset{(a)}{=} \sum_{i=1}^{n} \Big[I(W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1};Y_{3,i}) + I(W_1;Y_{2,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_2^{i+1})\\
&\quad +I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})\Big] + n(\gamma_3+\gamma_5+\gamma_8)\\
&= \sum_{i=1}^{n} [I(U_{3,i};Y_{3,i}) + I(U_{2,i};Y_{2,i}|U_{1,i}) + I(X_i;Y_{1,i}|U_{2,i},U_{3,i})] + n(\gamma_3+\gamma_5+\gamma_8), \quad (131)
\end{aligned}$$

where (a) is due to $I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) = I(\tilde{\mathbf{Y}}_2^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})$ by Relation 1.

For Case 3, we have

$$\begin{aligned}
n(R_0+R_1+R_2) &= H(W_0,W_1) + H(W_2|W_0,W_1)\\
&= I(W_0,W_1;\mathbf{Y}_2) + H(W_0,W_1|\mathbf{Y}_2) + I(W_2;\mathbf{Y}_1|W_0,W_1) + H(W_2|W_0,W_1,\mathbf{Y}_1)\\
&\leq \sum_{i=1}^{n} \Big[I(W_0,W_1;Y_{2,i}|\tilde{\mathbf{Y}}_2^{i+1}) + I(W_2;Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1})\Big] + n(\gamma_5+\gamma_7)\\
&\overset{(a)}{\leq} \sum_{i=1}^{n} \Big[I(W_0,W_1;Y_{2,i}|\tilde{\mathbf{Y}}_2^{i+1}) + I(X_i;Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1})\Big] + n(\gamma_5+\gamma_7) \quad (132)\\
&\overset{(b)}{\leq} \sum_{i=1}^{n} [I(U_{2,i};Y_{2,i}) + I(X_i;Y_{1,i}|U_{2,i},U_{3,i})] + n(\gamma_5+\gamma_7), \quad (133)
\end{aligned}$$

where (a) is by $W_2 \to X_i \to Y_{1,i}$; and (b) is by following the steps in (118) for the first term in the sum of (132) and the steps from (119)-(120) for the second term in the sum of (132).

Finally, introduce random variable G, which is independent of all other random variables and taking on values i, for $i = 1,2,\ldots,n$, with probability $1/n$. Define $U_k \triangleq (G,U_{k,G})$, $X \triangleq X_G$, $Y_k \triangleq Y_{k,G}$, $k = 1,2,3$. Then, we can obtain the rate region in Theorem 2 using (94), (99), (104), (108), (110), (112), (113), (114), (116), (117), (118), (122), (123), (126), (127), (130), (131) and (133).

B. Proof of the outer bound for the 3-receiver BC with 2 degraded message sets (Type 1)

In this section we show the proof for the outer bound of Corollary 2. The same code construction as in Section V-A, and preserve the definitions for the auxiliary random variables.

We begin with the equivocation rate R_{1e}. Following the same procedure to obtain (89) - (91), we have

$$
\begin{aligned}
nR_{1e} &\leq \sum_{i=1}^{n} \Big[I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1;Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1}) \Big] + n(\epsilon_1' + \gamma_3 + \gamma_4) \\
&= \sum_{i=1}^{n} \Big[I(W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1,\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1}) \\
&\quad -I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) + I(\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,W_1,\tilde{\mathbf{Y}}_3^{i+1}) \Big] + n(\epsilon_1' + \gamma_3 + \gamma_4) \\
&\overset{(a)}{=} \sum_{i=1}^{n} \Big[I(W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1,\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1}) \\
&\quad -I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) + I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) \Big] + n(\epsilon_1' + \gamma_3 + \gamma_4) \\
&= \sum_{i=1}^{n} \Big[I(W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(W_1,\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1}) \Big] + n(\epsilon_1' + \gamma_3 + \gamma_4) \\
&= \sum_{i=1}^{n} \Big[I(X_i,W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(X_i,W_1,\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1}) \\
&\quad - \Big(I(X_i;Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) - I(X_i;Y_{3,i}|W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) \Big) \Big] + n(\epsilon_1' + \gamma_3 + \gamma_4) \\
&\overset{(b)}{\leq} \sum_{i=1}^{n} \Big[I(X_i,W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(X_i,W_1,\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1}) \Big] + n(\epsilon_1' + \gamma_3 + \gamma_4) \\
&\overset{(c)}{=} \sum_{i=1}^{n} \Big[I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1}) - I(X_i;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) \Big] + n(\epsilon_1' + \gamma_3 + \gamma_4) \\
&= \sum_{i=1}^{n} \Big[I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1}) - I(X_i,\tilde{\mathbf{Y}}_3^{i+1};Y_{3,i}|W_0,\mathbf{Y}_1^{i-1}) \\
&\quad +I(\tilde{\mathbf{Y}}_3^{i+1};Y_{3,i}|W_0,\mathbf{Y}_1^{i-1}) \Big] + n(\epsilon_1' + \gamma_3 + \gamma_4) \\
&\overset{(d)}{=} \sum_{i=1}^{n} \Big[I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(X_i;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1}) \\
&\quad +I(\tilde{\mathbf{Y}}_3^{i+1};Y_{3,i}|W_0,\mathbf{Y}_1^{i-1}) \Big] + n(\epsilon_1' + \gamma_3 + \gamma_4) \\
&\overset{(e)}{=} \sum_{i=1}^{n} \Big[I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(\tilde{\mathbf{Y}}_2^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) - I(X_i;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1}) \\
&\quad +I(\tilde{\mathbf{Y}}_2^{i+1};Y_{3,i}|W_0,\mathbf{Y}_1^{i-1}) \Big] + n(\epsilon_1' + \gamma_3 + \gamma_4) \\
&= \sum_{i=1}^{n} \Big[I(X_i;Y_{1,i}|U_{1,i}) - I(\tilde{U}_{2,i};Y_{1,i}|U_{1,i}) - I(X_i;Y_{3,i}|U_{1,i}) + I(\tilde{U}_{2,i};Y_{3,i}|U_{1,i}) \Big] + n(\epsilon_1' + \gamma_3 + \gamma_4) \\
&\overset{(f)}{=} \sum_{i=1}^{n} \left[I(X_i;Y_{1,i}|U_{1,i}) - I(X_i;Y_{3,i}|U_{1,i}) \right] + n(\epsilon_1' + \gamma_3 + \gamma_4)
\end{aligned} \tag{134}
$$

where (a) is due to [1, Lemma 7] which gives

$$\sum_{i=1}^{n} I(\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,W_1,\tilde{\mathbf{Y}}_3^{i+1}) = \sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1});$$

(b) is due to the fact that Y_1 is a more capable channel than Y_3 so that $I(X_i;Y_{1,i}|W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) - I(X_i;Y_{3,i}|W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) \geq 0$ and is true as $(W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1}) \to X_i \to (Y_{1,i},Y_{3,i})$ forms a Markov chain so that the more capable channel condition is satisfied; (c) is because $I(W_1,\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1},X_i) = 0$ since given X_i, W_1 and $\tilde{\mathbf{Y}}_3^{i+1}$ are both independent of $Y_{1,i}$ from a functional dependency graph, and we also have $I(W_1;Y_{3,i}|W_0,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_3^{i+1},X_i) = 0$ since we have $W_1 \to X_i \to Y_{1,i}$; (d) has second term in the sum by [1, Lemma 7] by which we have

$$\sum_{i=1}^{n} I(\mathbf{Y}_1^{i-1};Y_{3,i}|W_0,\tilde{\mathbf{Y}}_3^{i+1}) = \sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}),$$

and third term by $I(\tilde{\mathbf{Y}}_3^{i+1};Y_{3,i}|W_0,\mathbf{Y}_1^{i-1},X_i) = 0$ since given X_i, $\tilde{\mathbf{Y}}_3^{i+1}$ is independent of $Y_{3,i}$; (e) is by Relation 1; and (f) is by substituting $\tilde{U}_{2,i} = U_{1,i}$.

For rates R_0 we already have, from (110), (112) the rates for W_0 sent to Y_1 and Y_3. For W_0 sent to Y_2, we have

$$\begin{aligned}
nR_0 &\leq \sum_{i=1}^{n} I(W_0;Y_{2,i}|\tilde{\mathbf{Y}}_2^{i+1}) + n\gamma_2 \\
&= \sum_{i=1}^{n} \Big[I(W_0,\mathbf{Y}_1^{i-1};Y_{2,i}|\tilde{\mathbf{Y}}_2^{i+1}) - I(\mathbf{Y}_1^{i-1};Y_{2,i}|W_0,\tilde{\mathbf{Y}}_2^{i+1})\Big] + n\gamma_2 \\
&\overset{(a)}{\leq} \sum_{i=1}^{n} \Big[I(W_0,W_1,\mathbf{Y}_1^{i-1},\tilde{\mathbf{Y}}_2^{i+1};Y_{2,i}) - I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})\Big] + n\gamma_2 \\
&= \sum_{i=1}^{n} \Big[I(W_0,W_1,\mathbf{Y}_1^{i-1};Y_{2,i}) + I(\tilde{\mathbf{Y}}_2^{i+1};Y_{2,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) - I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})\Big] + n\gamma_2 \\
&\overset{(b)}{=} \sum_{i=1}^{n} \Big[I(W_0,W_1,\mathbf{Y}_1^{i-1};Y_{2,i}) - I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1})\Big] + n\gamma_2 \\
&= \sum_{i=1}^{n} [I(U_{2,i},U_{1,i};Y_{2,i}) - I(U_{2,i};Y_{1,i}|U_{1,i})] + n\gamma_2, \qquad (135)
\end{aligned}$$

where (a) is due to

$$\begin{aligned}
\sum_{i=1}^{n} I(\mathbf{Y}_1^{i-1};Y_{2,i}|W_0,\tilde{\mathbf{Y}}_2^{i+1}) &= \sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_2^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) \\
= \sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_3^{i+1};Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) &= \sum_{i=1}^{n} I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) \qquad (136)
\end{aligned}$$

with the first equality due to [1, Lemma 7], the second and third equalities by Relations 1 and 2, respectively; (b) is due to $I(\tilde{\mathbf{Y}}_2^{i+1};Y_{2,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) = I(\tilde{\mathbf{Y}}_3^{i+1};Y_{2,i}|W_0,W_1,\mathbf{Y}_1^{i-1}) = 0$ by Relation 3 and (85b).

So we have

$$nR_0 \leq \sum_{i=1}^{n} I(U_{1,i};Y_{1,i}) + n\gamma_1, \tag{137}$$

$$nR_0 \leq \sum_{i=1}^{n} [I(U_{2,i};Y_{2,i}) - I(U_{2,i};Y_{1,i}|U_{1,i})] + n\gamma_2, \tag{138}$$

$$nR_0 \leq \sum_{i=1}^{n} [I(U_{3,i};Y_{3,i}) - I(U_{3,i};Y_{1,i}|U_{1,i})] + n\gamma_3. \tag{139}$$

For rates $R_0 + R_1$, consider W_0 sent to Y_1, Y_2, Y_3 and W_1 to Y_1 only. We have

$$\begin{aligned} n(R_0 + R_1) &\leq H(W_0) + \sum_{i=1}^{n} I(W_1;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) + n\gamma_4 \\ &\overset{(a)}{\leq} H(W_0) + \sum_{i=1}^{n} I(X_i;Y_{1,i}|W_0,\mathbf{Y}_1^{i-1}) + n\gamma_4, \end{aligned} \tag{140}$$

where (a) is by $W_1 \to X_i \to Y_{1,i}$. Then combining (140) with (137), (138), (139), respectively, we can get

$$n(R_0 + R_1) \leq \sum_{i=1}^{n} I(X_i;Y_{1,i}) + n(\gamma_1 + \gamma_4) \tag{141}$$

$$n(R_0 + R_1) \leq \sum_{i=1}^{n} [I(U_{2,i};Y_{2,i}) + I(X_i;Y_{1,i}|U_{2,i})] + n(\gamma_2 + \gamma_4) \tag{142}$$

$$n(R_0 + R_1) \leq \sum_{i=1}^{n} [I(U_{3,i};Y_{3,i}) + I(X_i;Y_{1,i}|U_{3,i})] + n(\gamma_3 + \gamma_4). \tag{143}$$

Now introduce the random variables G, X, Y_k, $k = 1,2,3$, and U_k, $k = 1,2$ as at the end of Section V-A, and using (134), (137), (138), (139), (141), (142) and (143), we can obtain the rate region in Corollary 2. So we have shown that the 3 degraded message set outer bound can reduce to the 2 degraded message set (Type 1) outer bound, as we have used the same condition (Y_1 more capable than Y_3), auxiliary random variable definition and code construction so that (81) is satisfied.

C. Proof for the 3-receiver BC with 2 degraded message sets (Type 2) with both Y_1 and Y_2 less noisy than Y_3

In this section we show the converse proof for the bound in Corollary 3. We now use a $(2^{nR_0}, 2^{nR_1}, n)$-code with error probability $P_e^{(n)}$ and code construction so that we have the Markov chain condition (W_0, W_1) $\to \mathbf{X} \to (\mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}_3)$. Then, the probability distribution on $\mathcal{W}_0 \times \mathcal{W}_1 \times \mathcal{X}^n \times \mathcal{Y}_1^n \times \mathcal{Y}_2^n \times \mathcal{Y}_3^n$ is given by

$$p(w_0)p(w_1)p(\mathbf{x}|w_0,w_1)\prod_{i=1}^{n} p(y_{1i}, y_{2i}, y_{3i}|x_i). \tag{144}$$

We first note that from the definition of more capable and less noisy channels [15], when Y_1 is less noisy than Y_2 or Y_3, then it also follows that Y_1 is more capable than Y_2 or Y_3.

We now also define the new auxiliary random variable $U_i \triangleq (W_0, \mathbf{Y}_3^{i-1})$ satisfying the condition

$$U_i \to X_i \to (Y_{1,i}, Y_{2,i}, Y_{3,i}), \qquad \forall i. \tag{145}$$

To proceed with the proof, we begin with the equivocation rates. We will consider 2 cases: the first, where W_1 is sent to Y_1, the second where W_1 is sent to Y_2. For W_1 sent to Y_1, we have, following (89)

$$\begin{aligned} nR_{1e} &\leq H(W_1|\mathbf{Y}_3) + n\epsilon_1' \text{ (by secrecy condition)} \\ &\leq I(W_1;\mathbf{Y}_1|W_0) - I(W_1;\mathbf{Y}_3|W_0) + n(\epsilon_1' + \gamma_3 + \gamma_4). \end{aligned} \tag{146}$$

Then the first two terms of (146) can be bounded as

$$\begin{aligned} & I(W_1;\mathbf{Y}_1|W_0) - I(W_1;\mathbf{Y}_3|W_0) = \sum_{i=1}^{n}\Big[I(W_1;Y_{1,i}|W_0,\tilde{\mathbf{Y}}_1^{i+1}) - I(W_1;Y_{3,i}|W_0,\mathbf{Y}_3^{i-1})\Big] \\ &= \sum_{i=1}^{n}\Big[I(W_1,\mathbf{Y}_3^{i-1};Y_{1,i}|W_0,\tilde{\mathbf{Y}}_1^{i+1}) - I(W_1,\tilde{\mathbf{Y}}_1^{i+1};Y_{3,i}|W_0,\mathbf{Y}_3^{i-1}) - I(\mathbf{Y}_3^{i-1};Y_{1,i}|W_0,W_1,\tilde{\mathbf{Y}}_1^{i+1}) \\ &\quad + I(\tilde{\mathbf{Y}}_1^{i+1};Y_{3,i}|W_0,W_1,\mathbf{Y}_3^{i-1})\Big] \\ &\overset{(a)}{=} \sum_{i=1}^{n}\Big[I(W_1,\mathbf{Y}_3^{i-1};Y_{1,i}|W_0,\tilde{\mathbf{Y}}_1^{i+1}) - I(W_1,\tilde{\mathbf{Y}}_1^{i+1};Y_{3,i}|W_0,\mathbf{Y}_3^{i-1}) - I(\tilde{\mathbf{Y}}_1^{i+1};Y_{3,i}|W_0,W_1,\mathbf{Y}_3^{i-1}) \\ &\quad + I(\tilde{\mathbf{Y}}_1^{i+1};Y_{3,i}|W_0,W_1,\mathbf{Y}_3^{i-1})\Big] \\ &= \sum_{i=1}^{n}\Big[I(W_1,\mathbf{Y}_3^{i-1};Y_{1,i}|W_0,\tilde{\mathbf{Y}}_1^{i+1}) - I(W_1,\tilde{\mathbf{Y}}_1^{i+1};Y_{3,i}|W_0,\mathbf{Y}_3^{i-1})\Big] \\ &= \sum_{i=1}^{n}\Big[I(X_i,W_1,\mathbf{Y}_3^{i-1};Y_{1,i}|W_0,\tilde{\mathbf{Y}}_1^{i+1}) - I(X_i,W_1,\tilde{\mathbf{Y}}_1^{i+1};Y_{3,i}|W_0,\mathbf{Y}_3^{i-1}) \\ &\quad - \Big(I(X_i;Y_{1,i}|W_0,W_1,\tilde{\mathbf{Y}}_1^{i+1},\mathbf{Y}_3^{i-1}) - I(X_i;Y_{3,i}|W_0,W_1,\tilde{\mathbf{Y}}_1^{i+1},\mathbf{Y}_3^{i-1})\Big)\Big] \\ &\overset{(b)}{\leq} \sum_{i=1}^{n}\Big[I(X_i,W_1,\mathbf{Y}_3^{i-1};Y_{1,i}|W_0,\tilde{\mathbf{Y}}_1^{i+1}) - I(X_i,W_1,\tilde{\mathbf{Y}}_1^{i+1};Y_{3,i}|W_0,\mathbf{Y}_3^{i-1})\Big] \\ &\overset{(c)}{=} \sum_{i=1}^{n}\Big[I(\mathbf{Y}_3^{i-1};Y_{1,i}|W_0,\tilde{\mathbf{Y}}_1^{i+1}) + I(X_i;Y_{1,i}|W_0,\tilde{\mathbf{Y}}_1^{i+1},\mathbf{Y}_3^{i-1}) - I(X_i;Y_{3,i}|W_0,\mathbf{Y}_3^{i-1})\Big] \\ &\overset{(d)}{=} \sum_{i=1}^{n}\Big[I(\tilde{\mathbf{Y}}_1^{i+1};Y_{3,i}|W_0,\mathbf{Y}_3^{i-1}) + I(X_i,\tilde{\mathbf{Y}}_1^{i+1};Y_{1,i}|W_0,\mathbf{Y}_3^{i-1}) - I(\tilde{\mathbf{Y}}_1^{i+1};Y_{1,i}|W_0,\mathbf{Y}_3^{i-1}) \\ &\quad - I(X_i;Y_{3,i}|W_0,\mathbf{Y}_3^{i-1})\Big] \\ &\overset{(e)}{=} \sum_{i=1}^{n}\Big[I(X_i;Y_{1,i}|W_0,\mathbf{Y}_3^{i-1}) - I(X_i;Y_{3,i}|W_0,\mathbf{Y}_3^{i-1}) - \Big(I(\tilde{\mathbf{Y}}_1^{i+1};Y_{1,i}|W_0,\mathbf{Y}_3^{i-1}) - I(\tilde{\mathbf{Y}}_1^{i+1};Y_{3,i}|W_0,\mathbf{Y}_3^{i-1})\Big)\Big] \\ &\overset{(f)}{\leq} \sum_{i=1}^{n}\Big[I(X_i;Y_{1,i}|W_0,\mathbf{Y}_3^{i-1}) - I(X_i;Y_{3,i}|W_0,\mathbf{Y}_3^{i-1})\Big] \end{aligned} \tag{147}$$

where (a) is by [1, Lemma 7] from which

$$\sum_{i=1}^{n} I(\mathbf{Y}_3^{i-1};Y_{1,i}|W_0,W_1,\tilde{\mathbf{Y}}_1^{i+1}) = \sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_1^{i+1};Y_{3,i}|W_0,W_1,\mathbf{Y}_3^{i-1});$$

(b) is by $I(X_i;Y_{1,i}|W_0,W_1,\tilde{\mathbf{Y}}_1^{i+1},\mathbf{Y}_3^{i-1}) - I(X_i;Y_{3,i}|W_0,W_1,\tilde{\mathbf{Y}}_1^{i+1},\mathbf{Y}_3^{i-1}) \geq 0$ as Y_1 is more capable than Y_3 which is a consequence of the assumption that Y_1 is less noisy than Y_3, with $(W_0,W_1,\tilde{\mathbf{Y}}_1^{i+1},\mathbf{Y}_3^{i-1}) \rightarrow X_i \rightarrow$

$(Y_{1,i}, Y_{3,i})$ fulfilling the more capable channel condition; (c) is because we have $I(W_1; Y_{1,i}|W_0, \tilde{\mathbf{Y}}_1^{i+1}, \mathbf{Y}_3^{i-1}, X_i) = 0$ since W_1 is independent of $Y_{1,i}$ given X_i and $I(W_1, \tilde{\mathbf{Y}}_1^{i+1}; Y_{3,i}|W_0, \mathbf{Y}_3^{i-1}, X_i) = 0$ since W_1 and $\tilde{\mathbf{Y}}_1^{i+1}$ are both independent of $Y_{3,i}$ given X_i, both of which can be verified using a functional dependency graph; (d) has first term by [1, Lemma 7] from which

$$\sum_{i=1}^{n} I(\mathbf{Y}_3^{i-1}; Y_{1,i}|W_0, \tilde{\mathbf{Y}}_1^{i+1}) = \sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_1^{i+1}; Y_{3,i}|W_0, \mathbf{Y}_3^{i-1});$$

(e) is by $I(\tilde{\mathbf{Y}}_1^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_3^{i-1}, X_i) = 0$ since $\tilde{\mathbf{Y}}_1^{i+1}$ is independent of $Y_{1,i}$ given X_i from a functional dependency graph; and (f) is due to $I(\tilde{\mathbf{Y}}_1^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_3^{i-1}) - I(\tilde{\mathbf{Y}}_1^{i+1}; Y_{3,i}|W_0, \mathbf{Y}_3^{i-1}) \geq 0$ from the fact that Y_1 is less noisy than Y_3. Thus we have

$$nR_{1e} \leq \sum_{i=1}^{n} [I(X_i; Y_{1,i}|U_i) - I(X_i; Y_{3,i}|U_i)] + n(\epsilon_1' + \gamma_3 + \gamma_4). \tag{148}$$

For rate R_{1e} arising from W_1 sent to Y_2, we follow the same procedure as in (146) to (148), except that all terms involving Y_1 are replaced with the corresponding terms involving Y_2, and carry out the expansion $I(W_1; \mathbf{Y}_2|W_0) = \sum_{i=1}^{n} I(W_1; Y_{2,i}|W_0, \tilde{\mathbf{Y}}_1^{i+1})$ instead, and the condition that Y_2 is less noisy than Y_3 is used. Then we can get

$$nR_{1e} \leq \sum_{i=1}^{n} [I(X_i; Y_{2,i}|U_i) - I(X_i; Y_{3,i}|U_i)] + n(\epsilon_1' + \gamma_2 + \gamma_4). \tag{149}$$

The rate R_0 may be easily found as

$$\begin{aligned} nR_0 = H(W_0) &\leq \sum_{i=1}^{n} I(W_0; Y_{3,i}|\mathbf{Y}_3^{i-1}) + n\gamma_3 \\ &\leq \sum_{i=1}^{n} I(W_0, \mathbf{Y}_3^{i-1}; Y_{3,i}) + n\gamma_3 = \sum_{i=1}^{n} I(U_i; Y_{3,i}) + n\gamma_3. \end{aligned} \tag{150}$$

For rates $R_0 + R_1$, first consider W_1 sent to receiver Y_1. We have

$$\begin{aligned} & n(R_0 + R_1) = H(W_0) + H(W_1|W_0) \leq H(W_0) + I(W_1; \mathbf{Y_1}|W_0) + n\gamma_4 \\ & = H(W_0) + \sum_{i=1}^{n} I(W_1; Y_{1,i}|W_0, \tilde{\mathbf{Y}}_1^{i+1}) + n\gamma_4 \overset{(a)}{\leq} H(W_0) + \sum_{i=1}^{n} I(X_i; Y_{1,i}|W_0, \tilde{\mathbf{Y}}_1^{i+1}) + n\gamma_4 \\ & \leq H(W_0) + \sum_{i=1}^{n} I(X_i, \mathbf{Y}_3^{i-1}; Y_{1,i}|W_0, \tilde{\mathbf{Y}}_1^{i+1}) + n\gamma_4 \\ & = H(W_0) + \sum_{i=1}^{n} \Big[I(\mathbf{Y}_3^{i-1}; Y_{1,i}|W_0, \tilde{\mathbf{Y}}_1^{i+1}) + I(X_i; Y_{1,i}|W_0, \tilde{\mathbf{Y}}_1^{i+1}, \mathbf{Y}_3^{i-1}) \Big] + n\gamma_4 \\ & \overset{(b)}{=} H(W_0) + \sum_{i=1}^{n} \Big[I(\tilde{\mathbf{Y}}_1^{i+1}; Y_{3,i}|W_0, \mathbf{Y}_3^{i-1}) + I(X_i, \tilde{\mathbf{Y}}_1^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_3^{i-1}) - I(\tilde{\mathbf{Y}}_1^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_3^{i-1}) \Big] + n\gamma_4 \\ & \overset{(c)}{\leq} H(W_0) + \sum_{i=1}^{n} I(X_i; Y_{1,i}|W_0, \mathbf{Y}_3^{i-1}) + n\gamma_4 \end{aligned} \tag{151}$$

where (a) is by $W_1 \to X_i \to Y_{1,i}$; (b) is by [1, Lemma 7] from which

$$\sum_{i=1}^{n} I(\mathbf{Y}_3^{i-1}; Y_{1,i}|W_0, \tilde{\mathbf{Y}}_1^{i+1}) = \sum_{i=1}^{n} I(\tilde{\mathbf{Y}}_1^{i+1}; Y_{3,i}|W_0, \mathbf{Y}_3^{i-1});$$

and (c) is because $I(\tilde{\mathbf{Y}}_1^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_3^{i-1}) - I(\tilde{\mathbf{Y}}_1^{i+1}; Y_{3,i}|W_0, \mathbf{Y}_3^{i-1}) \geq 0$ as Y_1 is less noisy than Y_3 and $I(\tilde{\mathbf{Y}}_1^{i+1}; Y_{1,i}|W_0, \mathbf{Y}_3^{i-1}, X_i) = 0$ as $\tilde{\mathbf{Y}}_1^{i+1}$ is independent of $Y_{1,i}$ given X_i from a functional dependency graph.

Next, for W_1 sent to Y_2, again follow the same procedure as to obtain (151), except that all terms involving Y_1 are replaced with the corresponding terms involving Y_2, and carry out the expansion $I(W_1; \mathbf{Y}_2|W_0) = \sum_{i=1}^{n} I(W_1; Y_{2,i}|W_0, \tilde{\mathbf{Y}}_1^{i+1})$, and the condition that Y_2 is less noisy than Y_3 is used. As such, we have

$$n(R_0 + R_1) \leq H(W_0) + \sum_{i=1}^{n} I(X_i; Y_{2,i}|W_0, \mathbf{Y}_3^{i-1}) + n\gamma_5. \tag{152}$$

For rates $(R_0 + R_1)$, considering (W_0, W_1) sent to receiver 1, we combine (150) with (151) to obtain

$$\begin{aligned} n(R_0 + R_1) &\leq \sum_{i=1}^{n} \Big[I(W_0, \mathbf{Y}_3^{i-1}; Y_{3,i}) + I(X_i; Y_{1,i}|W_0, \mathbf{Y}_3^{i-1}) \Big] + n(\gamma_3 + \gamma_4) \\ &= \sum_{i=1}^{n} \Big[I(W_0, \mathbf{Y}_3^{i-1}; Y_{1,i}) - \Big(I(W_0, \mathbf{Y}_3^{i-1}; Y_{1,i}) - I(W_0, \mathbf{Y}_3^{i-1}; Y_{3,i}) \Big) + I(X_i; Y_{1,i}|W_0, \mathbf{Y}_3^{i-1}) \Big] \\ &\quad + n(\gamma_3 + \gamma_4) \\ &\overset{(a)}{\leq} \sum_{i=1}^{n} I(X_i; Y_{1,i}) + n(\gamma_3 + \gamma_4) \end{aligned} \tag{153}$$

where (a) is by the condition $I(W_0, \mathbf{Y}_3^{i-1}; Y_{1,i}) - I(W_0, \mathbf{Y}_3^{i-1}; Y_{3,i}) \geq 0$ from Y_1 being less noisy than Y_3. Now considering (W_0, W_1) sent to receiver 2, combine (150) with (152) in the same way to obtain

$$n(R_0 + R_1) \leq \sum_{i=1}^{n} I(X_i; Y_{2,i}) + n(\gamma_3 + \gamma_5) \tag{154}$$

where now we have $I(W_0, \mathbf{Y}_3^{i-1}; Y_{2,i}) - I(W_0, \mathbf{Y}_3^{i-1}; Y_{3,i}) \geq 0$ from Y_2 being less noisy than Y_3.

Finally, introduce the random variables G, X, Y_k, $k = 1, 2, 3$, as at the end of Section V-A, and the random variable $U \triangleq (G, U_G)$. Using (150), (148), (149), (153) and (154), we obtain the rate region in Corollary 3. Thus we have shown that the outer bound for this 3-receiver, 2 degraded message set (Type 2) channel is a specialization of the more general 3-receiver, 3 degraded message set channel. We note that the outer bound to the rate equivocation region in Corollary 3 also coincides with a special case of the achievable bound of Chia and El Gamal [9] stated in Theorem 1 of [9], for the same message destinations and secrecy conditions.

VI. Conclusion

Bounds to the rate-equivocation region for the general 3-receiver BC with degraded message sets, in which receiver 3 is a wiretapper receiving the common message, are presented. This model is a more general model than the 2-receiver BCs with confidential messages with an external wiretapper, and 3-receiver degraded BCs with confidential messages. We obtain, with secrecy, new inner and outer bounds to the rate-equivocation region for the 3-receiver BC with 3 degraded message sets. We also obtain, without secrecy, new outer bounds to the rate region for the general 3-receiver BC with 3 degraded message sets. Lastly, we obtain new inner and outer bounds for rate-equivocation region for the 3-receiver BC with 2 degraded message sets (Type 1).

In the proof of achievability for the inner bound, we used Wyner's code partitioning combined with double-binning for secrecy. We have shown that the proposed coding scheme can provide security for the 3-receiver BC with 3 degraded message sets or 2 degraded message sets (Type 1), although the 2 degraded message set case (Type 1) will suffer a loss in the secrecy rate. The proof for the outer bound is shown for the 3-receiver BC with 3 degraded message sets and 2 degraded message sets (Type 1) under the condition that receiver 1 is more capable than receiver 3 the wiretapper; and for the 3-receiver BC with 2 degraded message sets (Type 2) for receivers 1 and 2 less noisy than the wiretapper. The outer bound for the 3 degraded message set case is shown to specialize to the 2 degraded message set (Type 1). Under the condition that both receivers 1 and 2 are less noisy than the wiretapper, the inner and outer bounds for the 3 degraded message case coincide and specialize to the rate-equivocation region of the 3-receiver BC with 2 degraded message sets (Type 2), and to a special case of a 3-receiver BC with 2 degraded message sets (Type 2) which uses a different coding scheme.

Appendix

Here, we show that we can insert an auxiliary random variable $\tilde{U}_2$, representing information about W_0, between U_1 and U_2, for the 3-receiver BC with 3 degraded message sets. We show that the conditions for correct code generation and low probability of error for decoding are equivalent to those without insertion of $\tilde{U}_2$ by setting $\tilde{U}_2 = U_1$. Thus, by their equivalence, we shall subsequently use the code generation process with the insertion of $\tilde{U}_2$ to facilitate the derivation of the outer bound.

We first note that such an insertion of $\tilde{U}_2$ gives rise to the Markov chains

$$U_1 \to \tilde{U}_2 \to U_2 \to (U_3, X) \to (Y_1, Y_2, Y_3), \tag{155a}$$

$$U_1 \to U_3 \to (\tilde{U}_2, U_2, X) \to (Y_1, Y_2, Y_3), \tag{155b}$$

$$U_1 \to (\tilde{U}_2, U_2, U_3) \to X \to (Y_1, Y_2, Y_3). \tag{155c}$$

Codebook generation is done as follows: first, generate 2^{nR_0} sequences $\mathbf{U}_1(w_0)$. Then, for each $\mathbf{U}_1(w_0)$, generate $2^{n\tilde{Q}_2}$ sequences $\tilde{\mathbf{U}}_2(w_0, \tilde{q}_2)$ and partition them into $2^{n\tilde{P}_2}$ equal-sized bins, and also 2^{nQ_3} sequences $\mathbf{U}_3(w_0, q_3)$. For each $\tilde{\mathbf{U}}_2(w_0, \tilde{p}_2)$, generate 2^{nQ_2} sequences $\mathbf{U}_2(w_0, \tilde{p}_2, q_2)$ and partition them into $2^{n\tilde{R}_1}$ bins. Also partition the $\mathbf{U}_3(w_0, q_3)$ into $2^{n\tilde{P}_3}$ equally-sized bins.

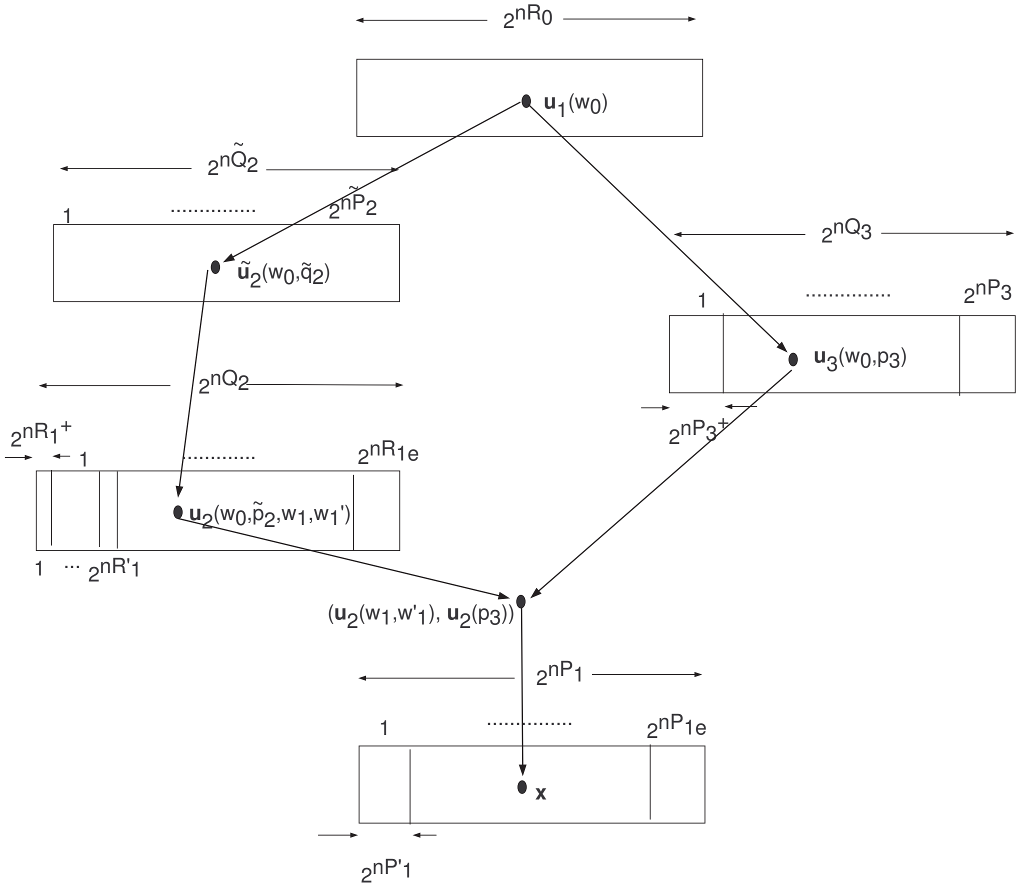


Fig. 5. Coding for 3-receiver BC with degraded message sets and confidential messages: insertion of auxiliary random variable $\tilde{U}_2$.

Each product bin (w_1, w_1', p_3) contains the joint typical pair $(\mathbf{U}_2(w_0, \tilde{p}_2, w_1, w_1', w_1^\dagger), \mathbf{U}_3(w_0, p_3, p_3^\dagger))$ with high probability under the conditions [16]

$$\begin{cases} \tilde{P}_2 \leq \tilde{Q}_2, \\ R_{1e} + R_1' \leq Q_2, \\ P_3 \leq Q_3, \\ \tilde{P}_2 + P_3 \leq \tilde{Q}_2 + Q_3 - I(\tilde{U}_2; U_3|U_1), \\ \tilde{P}_2 + R_{1e} + R_1' + P_3 \leq \tilde{Q}_2 + Q_2 + Q_3 - I(U_2; U_3|U_1). \end{cases} \tag{156}$$

For the joint typical pair $(\mathbf{U}_2(w_0, \tilde{p}_2, w_1, w_1'), \mathbf{U}_3(w_0, p_3))$ corresponding to the product bin (w_1, w_1', p_3), generate 2^{nP_1} sequences of codewords $\mathbf{X}(w_0, \tilde{p}_2, w_1, w_1', p_3, p_1, p_1')$. The decoding follows from what described in Section IV-A. Assume that $(w_0, \tilde{p}_2, w_1, p_3, p_1) = (1, 1, 1, 1, 1)$ is sent. For receiver Y_1, joint typical decoding of $\{\mathbf{u}_1, \tilde{\mathbf{u}}_2, \mathbf{u}_2, \mathbf{u}_3, \mathbf{y}_1\}$ is carried out. We list the error events and the conditions that ensure low error probability when decoding, while noting that the decoding of $\tilde{p}_2$ and w_1 is independent:

1) $\Pr\{\mathtt{E}_1 : (w_0 \neq 1)\} \leq \epsilon$ when

$$R_0 + \tilde{P}_2 + R_{1e} + R_1' + P_{1e} + P_1' + P_3 < I(X;Y_1). \tag{157}$$

2) $\Pr\{\mathtt{E}_2 : (w_0 = 1, \tilde{p}_2 \neq 1)\} \leq \epsilon$ when

$$\tilde{P}_2 + R_{1e} + R_1' + P_{1e} + P_1' + P_3 < I(X;Y_1|U_1). \tag{158}$$

3) $\Pr\{\mathtt{E}_3 : (w_0 = 1, \tilde{p}_2 \neq 1, w_1 \neq 1)\} \leq \epsilon$ when

$$\tilde{P}_2 + R_{1e} + R_1' + P_{1e} + P_1' + P_3 < I(X;Y_1|U_1). \tag{159}$$

4) $\Pr\{\mathtt{E}_4 : (w_0 = 1, \tilde{p}_2 \neq 1, w_1 \neq 1, p_3 = 1)\} \leq \epsilon$ when

$$\tilde{P}_2 + R_{1e} + R_1' + P_{1e} + P_1' < I(X;Y_1|U_3,U_1) = I(X;Y_1|U_3). \tag{160}$$

5) $\Pr\{\mathtt{E}_5 : (w_0 = 1, \tilde{p}_2 \neq 1, w_1 = 1, p_3 \neq 1)\} \leq \epsilon$ when

$$\tilde{P}_2 + P_{1e} + P_1' + P_3 < I(X;Y_1|U_2,U_1) = I(X;Y_1|U_2). \tag{161}$$

6) $\Pr\{\mathtt{E}_6 : (w_0 = 1, \tilde{p}_2 \neq 1, w_1 = 1, p_3 = 1, p_1 \neq 1)\} \leq \epsilon$ when

$$\tilde{P}_2 + P_{1e} + P_1' < I(X;Y_1|U_2,U_3,U_1) = I(X;Y_1|U_2,U_3). \tag{162}$$

7) $\Pr\{\mathtt{E}_7 : (w_0 = 1, \tilde{p}_2 = 1, w_1 \neq 1)\} \leq \epsilon$ when

$$\tilde{R}_{1e} + R_1' + P_{1e} + P_1' + P_3 < I(X;Y_1|\tilde{U}_2,U_1) = I(X;Y_1|\tilde{U}_2). \tag{163}$$

8) $\Pr\{\mathtt{E}_8 : (w_0 = 1, \tilde{p}_2 = 1, w_1 \neq 1, p_3 = 1)\} \leq \epsilon$ when

$$\tilde{R}_{1e} + R_1' + P_{1e} + P_1' < I(X;Y_1|U_3,\tilde{U}_2,U_1) = I(X;Y_1|U_3,\tilde{U}_2). \tag{164}$$

9) $\Pr\{\mathtt{E}_9 : (w_0 = 1, \tilde{p}_2 = 1, w_1 = 1, p_3 \neq 1)\} \leq \epsilon$ when

$$\tilde{P}_{1e} + P_1' + P_3 < I(X;Y_1|\tilde{U}_2,U_2) = I(X;Y_1|U_2). \tag{165}$$

10) $\Pr\{\mathtt{E}_1 0 : (w_0 = 1, \tilde{p}_2 = 1, w_1 = 1, p_3 = 1, p_1 \neq 1)\} \leq \epsilon$ when

$$\tilde{P}_{1e} + P_1' < I(X;Y_1|\tilde{U}_2,U_2,U_3,U_1) = I(X;Y_1|U_2,U_3). \tag{166}$$

Receiver Y_2 finds $(w_0, \tilde{q}_2)$ by indirectly decoding U_2, and w_1 by decoding U_2 conditioned on $(\tilde{U}_2, U_1)$. As a result, we have the conditions

$$R_0 + \tilde{Q}_2 + Q_2 < I(U_2;Y_2), \tag{167}$$

$$Q_2 < I(U_2;Y_2|\tilde{U}_2,U_1) = I(U_2;Y_2|\tilde{U}_2). \tag{168}$$

Receiver Y_3 finds w_0 by indirectly decoding U_3, which has low probability of error under the condition

$$R_0 + Q_3 < I(U_3; Y_3). \tag{169}$$

Compare the above conditions with the conditions for the 3-receiver BC with 3 degraded message sets without insertion of $\tilde{U}_2$ found in (18), (21), (24), (26), (28), (30), (32), (34) and (35). By setting $\tilde{U}_2 = U_1$, the conditions (156), (157)–(169) are maximized. Furthermore, by setting $\tilde{P}_2 = \tilde{Q}_2 = 0$, the conditions (156), (157)–(169) are equivalent to those in (18)–(35). Thus, we may insert $\tilde{U}_2$ representing information about W_0 between U_1 and U_2 giving the Markov chain conditions (155), and the conditions on decoding and code generation thus obtained are equivalent to the original conditions with $\tilde{U}_2 = U_1$. As such, we can derive the outer bound in 2 steps. In the first step, we insert $\tilde{U}_2$ and use Markov chain conditions (155) to obtain an outer bound $\mathcal{R}'_O$ which is equivalent to the one with original conditions (9) by setting $\tilde{U}_2 = U_1$. Then, set $\tilde{U}_2 = U_1$ in $\mathcal{R}'_O$ to obtain $\mathcal{R}_O$.

REFERENCES

[1] I. Csiszár and J. Körner, “Broadcast channels with confidential messages,” *IEEE Trans. Info. Theory*, vol. 24, no. 3, pp. 339–348, Mar. 1978.

[2] A. D. Wyner, “The wire-tap channel,” *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, 1975.

[3] R. Liu, I. Marić, P. Spasojević and R. Yates, “Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions,” *IEEE Trans. Info. Theory*, vol. 54, no. 6, Jun. 2008.

[4] J. Xu, Y. Cao, and B. Chen, “Capacity bounds for broadcast channels with confidential messages,” *IEEE Trans. Info. Theory*, vol. 55, no. 10, pp. 4529–4542, Oct. 2009.

[5] G. Bagherikaram, A. S. Motahari, and A. K. Khandani, “Secrecy rate region of the broadcast channel,” *IEEE Trans. Info. Theory*, submitted for publication, Jul. 2008.

[6] L. C. Choo and K. K. Wong, “The K-receiver broadcast channel with confidential messages,” submitted to *IEEE Trans. Info. Theory*, Dec. 2008.

[7] E. Ekrem and S. Ulukus, “Secrecy capacity of a class of broadcast channels with an eavesdropper,” *EURASIP J. Wireless Commun. and Net., Special issue on Wireless Physical Layer Security*, June 2009.

[8] M. Kobayashi, M. Debbah, and S. Shamai, “Secured communication over frequency-selective fading channels: A practical Vandermonde precoding,” *EURASIP J. Wireless Commun. and Net., Special issue on Wireless Physical Layer Security*, June 2009.

[9] Y.-K. Chia and A. El Gamal, “3-receiver broadcast channels with common and confidential messages,” *IEEE Int. Symp. Info. Theory 2009*, June 28–July 3, 2009.

[10] C. Nair and A. El Gamal, “The capacity region of a class of 3-receiver broadcast channels with degraded message sets,” *IEEE Int. Symp. Info. Theory 2008*, Toronto, July 6 –11, 2008.

[11] C. Nair and A. El Gamal, “The capacity region of a class of 3-receiver broadcast channels with degraded message sets,” submitted to *IEEE Trans. Info. Theory*, Dec. 2007. [Online] Available: http://arxiv.org/abs/0712.3327

[12] C. Nair and A. El Gamal, “The capacity region of a class of three-receiver broadcast channels with degraded message sets,” *IEEE Trans. Info. Theory*, vol. 55, no. 10, pp. 4479–4493, Oct. 2009.

[13] K. Marton, “A coding scheme for the discrete memoryless broadcast channel,” *IEEE Trans. Info. Theory*, vol. 25, no. 3, pp. 306–311, 1979.

[14] L. C. Choo and K. K. Wong, “Physical layer security for a 3-receiver broadcast channel with degraded message sets,” to appear, *Int. Conf. Wireless Comms. and Signal Processing 2009*, Nov. 13–15, Nanjing, China, 2009.

[15] J. Körner and K. Marton, “Comparison of two noisy channels,” *Topics in Information Theory, Keszthely, Hungary, 1975, Colloquia Math. Soc. Janos Bolyai*, North-Holland, pp. 411–423, 1977.

[16] A. El Gamal and E.C. van der Meulen, “A proof of Marton’s coding theorem for the discrete memoryless broadcast channel,” *IEEE Trans. Info. Theory*, vol. 27, no. 1, pp. 120–122, Jan. 1981.

[17] G. Kramer, “Topics in multi-user information theory,” *Foundations and Trends in Commun. and Info. Theory*, vol. 4, no.s 4–5, pp. 265–444, 2007.